Example 1

Exponential Data set

The model fitted to the exponential data (handout #1) is of the form:

$$y = ae^{-b*x} \quad .$$

The required SAS codes for fitting this model are as follows:

```
PROC NLIN DATA=EXP METHOD=GAUSS;
     PARMS A=-156 B=.2;

     MODEL Y = A*EXP(-B*X);
     DER.A = EXP(-B*X);
     DER.B = -A*X*EXP(-B*X);
     OUTPUT OUT=PRED P=PR;

/************************************************/
/**       Plotting the data and predicted function      **/
/************************************************/
     symbol1 i=none v=diamond c=black;
     symbol2 i=join v=none c=orange;

     proc gplot;
             plot y*x=1 pr*x=2/overlay;
     run;
```

The resulting output contains several pieces of information. The first is the iteration history which lists all the values tried for the parameters with the resulting SSE. If a good estimate is found, a message indicating "convergence criterion met" will be printed. Next is the table of sum of squares and degrees of freedom followed by the parameters estimates and their asymptotic standard errors. Note that these values are linear approximations (based on a normality assumption) and may not be reliable (diagnostics to evaluate these estimates are given in handout #4). The last table printed is of parameter estimate correlations since nonlinear regression may result in dependency among the parameters. Very high correlations (>> .99) may indicate that the model form is inappropriate for the data (see handout #4).

```
          Non-Linear Least Squares Iterative Phase
        Dependent Variable Y   Method: Gauss-Newton
     Iter              A              B Sum of Squares
        0    -156.000000       0.200000   13397.981027
        1    -156.947570       0.199656   13390.093550
        2    -156.939560       0.199672   13390.093133
        3    -156.939941       0.199672   13390.093132
NOTE: Convergence criterion met.
```

```
     Non-Linear Least Squares Summary Statistics     Dependent Variable Y

       Source                    DF Sum of Squares      Mean Square


       Regression                 2    327028.84917     163514.42458
       Residual                   4     13390.09313       3347.52328
       Uncorrected Total          6    340418.94230

       (Corrected Total)          5    110005.33333

       Parameter      Estimate     Asymptotic            Asymptotic 95 %
                                   Std. Error         Confidence Interval
                                                      Lower          Upper
              A   -156.9399409  29.857773072 -239.83728713 -74.042594654
              B      0.1996716   0.046504675    0.07055566   0.328787521

                         Asymptotic Correlation Matrix

                 Corr                        A                 B
                 ----------------------------------------------
                        A                    1      0.8468726626
                        B        0.8468726626                 1
```

Here the iteration process converges in 3 steps resulting in parameter estimates which are asymptotically significant with reasonable correlation. It is often helpful to examine the iteration history. In this case, the estimates are close to the starting values and have not changed radically. This leads to a quick convergence indicating a potentially reasonable fit.

Example 2

Prickly Lettuce Data set

The model fitted to the prickly lettuce data is:

$$\mathbf{germ = M/(1 + e^{(-B(time-L))})}$$

which is a form of the logistic growth model. With the above model specification (parameterization): **M** represents the maximum attainable germination, **B** relates to the rate of germination, and **L** measures the time to 50% germination. It is important to note that each of these parameters has a relevant biological interpretation. Choice of the nonlinear model should not be arbitrary, but rather it should be selected carefully to parametrically represent the underlying biological process.

Codes and results for this data follow. Notice the use of the BY statement (BY processing) to fit both biotypes in one run.

Again, both data sets converge quickly to asymptotically significant estimates with low correlations. In fact, these fits are very impressive as indicated by the narrow range of the confidence intervals and the size of the regression sum of squares relative to the total sum of squares.

```
PROC NLIN DATA=plett METHOD=gauss;
    PARMS M=100 L=11 B=1;
    MODEL germ=M/(1+EXP(-B*(time-L)));
    DER.M =1/(1+EXP(-B*(time-L)));
    DER.L=-(M*(B*exp(-(B*(time-L))))/(1+exp(-(B*(time-L))))**2);
    DER.B=-(M*(-((time-L)*exp(-(B*(time-L)))))/
     (1+exp(-(B*(time-L))))**2);
    OUTPUT OUT=PRED1 P=PR student=RESID L95=low u95=up;
        by bio;

/************************************************/
/**      Plotting the data and predicted function       **/
/************************************************/

SYMBOL1 V=diamond I=NONE C=black;
SYMBOL2 V=none I=join w=2 C=orange;

proc gplot data=pred1 uniform;
    plot emerge*time=1 PR*time=2/overlay;
        by bio;
```

```
--------------------------------- BIO=R -------------------------------------

                    Non-Linear Least Squares Iterative Phase
                 Dependent Variable GERM   Method: Gauss-Newton
       Iter             M              L               B  Sum of Squares
          0    100.000000      11.000000        1.000000    7996.512456
          1     99.339808      11.418976        1.251612    5329.871233
          2     99.267290      11.363997        1.374471    5197.782077
          3     99.276442      11.367607        1.389194    5196.456011
          4     99.276056      11.367674        1.389950    5196.453173
          5     99.276037      11.367679        1.389981    5196.453168
NOTE: Convergence criterion met.
     Non-Linear Least Squares Summary Statistics     Dependent Variable GERM

        Source                DF Sum of Squares     Mean Square

        Regression             3   1680179.5468     560059.8489
        Residual             237      5196.4532         21.9260
        Uncorrected Total    240   1685376.0000

        (Corrected Total)    239    431115.5833


        Parameter      Estimate     Asymptotic            Asymptotic 95 %
                                    Std. Error         Confidence Interval
                                                       Lower          Upper
               M    99.27603715 0.37137504235  98.544410201  100.00766410
               L    11.36767872 0.04082659073  11.287248347   11.44810910
               B     1.38998060 0.06903338146   1.253981468    1.52597973
```

Asymptotic Correlation Matrix

Corr	M	L	B
M	1	0.198073651	-0.175030454
L	0.198073651	1	-0.035966444
B	-0.175030454	-0.035966444	1

--------------------------------- BIO=S ----------------------------------

Non-Linear Least Squares Iterative Phase

Dependent Variable GERM Method: Gauss-Newton

Iter	M	L	B	Sum of Squares
0	100.000000	11.000000	1.000000	131061
1	97.676506	12.280100	0.274031	68721.630874
2	100.471923	16.150336	0.414255	28652.194013
3	96.775367	14.391708	0.628096	13339.669054
4	98.367783	14.909415	0.868108	8950.259912
5	98.358988	14.767388	0.941069	8656.280566
6	98.331149	14.759881	0.957959	8651.679063
7	98.325992	14.757036	0.958119	8651.607276
8	98.325200	14.756971	0.958425	8651.606014
9	98.325116	14.756919	0.958417	8651.605991

NOTE: Convergence criterion met.

Non-Linear Least Squares Summary Statistics   Dependent Variable GERM

Source	DF	Sum of Squares	Mean Square
Regression	3	1432754.3940	477584.7980
Residual	237	8651.6060	36.5047
Uncorrected Total	240	1441406.0000	
(Corrected Total)	239	487081.1833	

Parameter	Estimate	Asymptotic Std. Error	Asymptotic 95 % Confidence Interval Lower	Upper
M	98.32511573	0.52682418615	97.287246386	99.362985072
L	14.75691897	0.06496914808	14.628926584	14.884911365
B	0.95841677	0.05215153887	0.855675694	1.061157844

Asymptotic Correlation Matrix

Corr	M	L	B
M	1	0.2581412342	-0.229013018
L	0.2581412342	1	-0.05910854
B	-0.229013018	-0.05910854	1