Fitting Nonlinear Mixed Models with the New NLMIXED Procedure

Russell D. Wolfinger, SAS Institute Inc., Cary, NC

ABSTRACT

Statistical models in which both fixed and random effects enter nonlinearly are becoming increasingly popular. These models have a wide variety of applications, two of the most common being nonlinear growth curves and overdispersed binomial data. A new SAS/STAT® procedure, NLMIXED, fits these models using likelihood-based methods. This paper presents some of the primary features of PROC NLMIXED and illustrates its use with two examples.

INTRODUCTION

The NLMIXED procedure fits nonlinear mixed models, that is, models in which both fixed and random effects are permitted to have a nonlinear relationship to the response variable. These models can take various forms, but the most common ones involve a conditional distribution for the response variable given the random effects. PROC NLMIXED enables you to specify such a distribution by using either a keyword for a standard form (normal, binomial, Poisson) or SAS programming statements to specify a general distribution.

PROC NLMIXED fits the specified nonlinear mixed model by maximizing an approximation to the likelihood integrated over the random effects. Different approximations to the integral are available, and the two principal ones are adaptive Gaussian quadrature and a first-order Taylor series approximation. You can use a variety of alternative optimization techniques to carry out the maximization; the default is a dual quasi-Newton algorithm.

Successful convergence of the optimization problem results in parameter estimates along with their approximate standard errors computed from the second derivative matrix of the likelihood function. PROC NLMIXED enables you to use the estimated model to construct predictions of arbitrary functions by using the parameter estimates and the empirical Bayes estimates of the random effects. PROC NLMIXED approximates their standard errors using the first derivatives of the function that you specify (the delta method).

LITERATURE ON NONLINEAR MIXED MODELS

Davidian and Giltinan (1995) and Vonesh and Chinchilli (1996) provide good overviews as well as general theoretical developments and examples of nonlinear mixed models. Pinheiro and Bates (1995) is a primary reference for the theory and computational techniques of PROC NLMIXED. They describe and compare several different integrated likelihood approximations and provide evidence that adaptive Gaussian quadrature is one of the best methods. Davidian and Gallant (1993) also use Gaussian quadrature for nonlinear mixed models, although the smooth nonparametric density they advocate for the random effects is currently not available in PROC NLMIXED.

Traditional approaches to fitting nonlinear mixed models involve Taylor series expansions, expanding around either zero or the empirical best linear unbiased predictions of the random effects. The former is the basis for the well-known first-order method of Beal and Sheiner (1982, 1988) and Sheiner and Beal (1985), and it is implemented in PROC NLMIXED. The latter is the basis for the estimation method of Lindstrom and Bates (1990), and it is not available in PROC NLMIXED. However, the closely related Laplacian approximation is available; it is equivalent to adaptive Gaussian quadrature with only one quadrature point. The Laplacian approximation and its relationship to the Lindstrom-Bates method are discussed by Beal and Sheiner (1992), Wolfinger (1993), Vonesh (1992, 1996), and Wolfinger and Lin (1997).

A parallel literature exists in the area of generalized linear mixed models, in which random effects appear as a part of the linear predictor inside of a link function. Taylor-series methods that are similar to those just described are discussed in articles such as Harville and Mee (1984), Stiratelli, Laird, and Ware (1984), Gilmour, Anderson, and Rae (1985), Goldstein (1991), Schall (1991), Engel and Keen (1992), Breslow and Clayton (1993), Wolfinger and O'Connell (1993), and McGilchrist (1994), but such methods have not been implemented in PROC NLMIXED because they can produce biased results in certain binary data situations (Rodriguez and Goldman 1995, Lin and Breslow 1996). Instead, a numerical quadrature approach is available in PROC NLMIXED, as discussed in Pierce and Sands (1975), Anderson and Aitkin (1985), Crouch and Spiegelman (1990), Hedeker and Gibbons (1994), Longford (1994), McCulloch (1994), Liu and Pierce (1994), and Diggle, Liang, and Zeger (1994).

PROC NLMIXED COMPARED WITH OTHER SAS PROCEDURES AND MACROS

The models fit by PROC NLMIXED can be viewed as generalizations of the random coefficient models fit by the MIXED procedure. This generalization allows the random coefficients to enter the model nonlinearly, whereas in PROC MIXED they enter linearly. Because of this general nonlinear formulation, no direct analog to the REML method is available in PROC NLMIXED; only standard maximum likelihood is used. Also, PROC MIXED assumes the data to be normally distributed, whereas PROC NLMIXED enables you to analyze data that are normal, binomial, or Poisson or that have any likelihood programmable with SAS statements.

PROC NLMIXED does not implement the same estimation techniques that are available with the NLINMIX and GLIMMIX macros. These macros are based on the estimation methods of Lindstrom and Bates (1990), Breslow and Clayton (1993), and Wolfinger and O'Connell (1993), and they iteratively fit a set of generalized estimating equations (refer to Chapters 11 and 12 of Littell et al. 1996 and to Wolfinger 1997). In contrast, PROC NLMIXED directly maximizes an approximate integrated likelihood.

PROC NLMIXED has close ties with the NLP procedure in SAS/OR® software. PROC NLMIXED uses a subset of the optimization code underlying PROC NLP and has many of the same optimization-based options. Also, the programming statement functionality that is used by PROC NLMIXED is the same as that used by PROC NLP and the MODEL procedure in SAS/ETS® software.

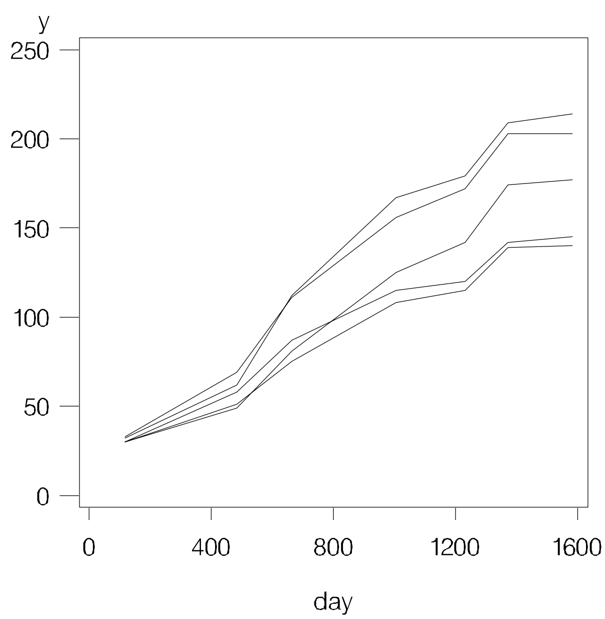
NONLINEAR GROWTH CURVES WITH GAUSSIAN DATA

As our first example, consider the orange tree data of Draper and Smith (1981). These data consist of seven measurements of the trunk circumference (in millimeters) on each of five orange trees. You can input these data into a SAS data set as follows:

```
data tree;
   input tree day y;
   datalines;
```

```
1  118   30
1  484   58
1  664   87
1 1004  115
1 1231  120
1 1372  142
1 1582  145
2  118   33
2  484   69
2  664  111
2 1004  156
2 1231  172
2 1372  203
2 1582  203
3  118   30
3  484   51
3  664   75
3 1004  108
3 1231  115
3 1372  139
3 1582  140
4  118   32
4  484   62
4  664  112
4 1004  167
4 1231  179
4 1372  209
4 1582  214
5  118   30
5  484   49
5  664   81
5 1004  125
5 1231  142
5 1372  174
5 1582  177
run;
```

The following is a plot of the profiles of the trees. Each profile has a flattish S shape, and between-tree variability increases with days.



Lindstrom and Bates (1990) and Pinheiro and Bates (1995) propose the following logistic nonlinear mixed model for these data:

$$y_{ij} = \frac{b_1 + u_{i1}}{1 + \exp[-(d_{ij} - b_2)/b_3]} + e_{ij}$$

Here, y_{ij} represents the jth measurement on the ith tree (i=1,…,5; j=1,…,7); d_{ij} is the corresponding day; b_1, b_2, b_3 are the fixed-effects parameters; u_{i1} are the random-effect parameters assumed to be iid $N(0,\sigma^2_u)$, and e_{ij} are the residual errors assumed to be iid $N(0,\sigma^2_e)$ and independent of the u_{i1}. This model has a logistic form, and the random-effect parameters u_{i1} enter the model linearly.

The PROC NLMIXED statements to fit this nonlinear mixed model are as follows:

```
proc nlmixed data=tree;
   parms b1=190 b2=700 b3=350 s2u=1000
      s2e=60;
   num = b1+u1;
   ex = exp(-(day-b2)/b3);
   den = 1 + ex;
   model y ~ normal(num/den,s2e);
   random u1 ~ normal(0,s2u) subject=tree;
run;
```

The PROC NLMIXED statement invokes the procedure and inputs the TREE data set. The PARMS statement identifies the unknown parameters and their starting values. Here, there are three fixed-effects parameters (B1, B2, B3) and two variance components (S2U, S2E).

The next three statements are SAS programming statements specifying the logistic mixed model. A new variable U1 is included to identify the random effect. These statements are evaluated for every observation in the data set when PROC NLMIXED computes the integrated log likelihood function and its derivatives.

The MODEL statement defines the dependent variable and its conditional distribution given the random effects. Here a normal (Gaussian) conditional distribution is specified with mean NUM/DEN and variance S2E.

The RANDOM statement defines the single random effect to be U1, and specifies that it follows a normal distribution with mean 0 and variance S2U. The SUBJECT= argument defines a variable indicating when the random effect obtains new realizations; in this case, it changes according to the values of the TREE variable. PROC NLMIXED assumes that the input data set is clustered according to the levels of the TREE variable; that is, all observations from the same tree occur sequentially in the input data set.

The output from this analysis is as follows.

The NLMIXED Procedure

Specifications

Data Set	WORK.TREE
Dependent Variable	y
Distribution for Dep Variable	Normal
Random Effects	u1
Distribution for Random Effects	Normal
Subject Variable	tree
Optimization Technique	Dual Quasi-Newton
Estimation Method	Adaptive Gaussian Quadrature

The "Specifications" table lists some basic information about the nonlinear mixed model that you have specified. Included are the input data set, dependent and subject variables, random effects, relevant distributions, and type of optimization.

Dimensions

Observations Used	35
Observations Not Used	0
Total Observations	35
Subjects	5
Max Obs Per Subject	7
Parameters	5
Quadrature Points	1

The "Dimensions" table lists various counts related to the model, including the number of observations, subjects, and parameters. These quantities are useful for checking that you have specified your data set and model correctly. Also listed is the number of quadrature points that PROC NLMIXED has selected based on the evaluation of the log likelihood at the starting values of the parameters. Here, only one quadrature point is necessary because the random-effect parameters enter the model linearly.

Parameters

b1	b2	b3	s2u	s2e	NegLogLike
190	700	350	1000	60	132.491787

The "Parameters" table lists the parameters to be estimated, their starting values, and the negative log likelihood evaluated at these starting values.

Iterations

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	4	131.686742	0.805045	0.010269	-0.633
2	6	131.6446	0.042082	0.014783	-0.0182
3	8	131.614077	0.030583	0.009809	-0.02796
4	10	131.57252	0.04155	0.001186	-0.01344
5	11	131.571895	0.000627	0.0002	-0.00121
6	13	131.571889	5.549E-6	0.00009	-7.68E-6
7	15	131.57188	1.096E-6	6.097E-6	-1.29E-6

NOTE: GCONV convergence criterion satisfied.

The "Iterations" table records the history of the minimization of the negative log likelihood. For each iteration of the quasi-Newton optimization, values are listed for the number of function calls, the value of the negative log likelihood, the difference from the previous iteration, the absolute value of the largest gradient, and the slope of the search direction. The note at the bottom of the table indicates that the algorithm has converged successfully according to the GCONV convergence criterion, a standard criterion computed using a quadratic form in the gradient and inverse Hessian.

```
              Fitting Information

   -2 Log Likelihood                     263.1
   AIC (smaller is better)               273.1
   BIC (smaller is better)               271.2
   Log Likelihood                       -131.6
   AIC (larger is better)               -136.6
   BIC (larger is better)               -135.6
```

The ”Fitting Information” table lists the final maximized value of the log likelihood as well as the information criteria of Akaike and Schwarz in two different forms. These statistics can be used to compare different nonlinear mixed models.

```
              Parameter Estimates

                  Standard         t
Parameter Estimate   Error   DF Value Pr > |t|

b1          192.05 15.6473   4 12.27   0.0003
b2          727.90 35.2472   4 20.65   <.0001
b3          348.07 27.0790   4 12.85   0.0002
s2u         999.88  647.44   4  1.54   0.1974
s2e        61.5139 15.8831   4  3.87   0.0179

Parameter  Alpha    Lower    Upper  Gradient

b1          0.05   148.61   235.50   1.154E-6
b2          0.05   630.04   825.76   5.289E-6
b3          0.05   272.88   423.25    -6.1E-6
s2u         0.05  -797.70  2797.45   -3.84E-6
s2e         0.05  17.4153   105.61   2.892E-6
```

The “Parameter Estimates” table lists the maximum likelihood estimates of the five parameters and their approximate standard errors computed using the final Hessian matrix. Approximate t values and Wald-type confidence limits are also provided, with degrees of freedom equal to the number of subjects minus the number of random effects. You should interpret these statistics cautiously for variance parameters like S2U and S2E because their sampling distributions tend to be skewed. The final column in the output is the gradient vector at the optimization solution. Each element appears to be sufficiently small to indicate a stationary point.

Since the random-effect parameters enter the model linearly, you can obtain equivalent results by using the first-order method (specify METHOD=FIRO in the PROC NLMIXED statement).

LOGISTIC-NORMAL MODEL WITH BINOMIAL DATA

Our second example concerns the data from Beitler and Landis (1985), which represent results from a multicenter clinical trial investigating the results of two topical cream treatments (active drug, control) in curing an infection. For each of eight clinics, the number of trials and favorable cures are recorded for each treatment. The SAS data set is as follows.

```
data infection;
   input clinic t x n;
   datalines;
1 1 11 36
1 0 10 37
2 1 16 20
2 0 22 32
3 1 14 19
3 0  7 19
4 1  2 16
4 0  1 17
5 1  6 17
5 0  0 12
6 1  1 11
6 0  0 10
7 1  1  5
7 0  1  9
8 1  4  6
8 0  6  7
run;
```

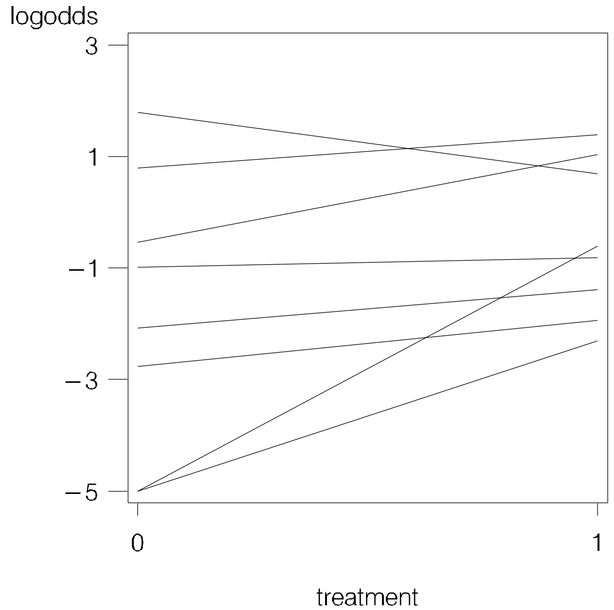
Suppose n_{ij} denotes the number of trials for the ith clinic and the jth treatment (i=1,…,8; j=0,1), and x_{ij} the corresponding number of favorable cures. Then a reasonable model for the preceding data is the following logistic model with random effects:

$$x_{ij} \mid u_i \sim \text{Binomial}(n_{ij}, p_{ij})$$

$$\eta_{ij} = \log\left(\frac{p_{ij}}{(1-p_{ij})}\right) = \beta_0 + \beta_1 t_j + u_i$$

The notation t_j indicates the jth treatment, and the u_i are assumed to be iid N(0, ,σ^2_u).

The observed log odds profile for each clinic is displayed in the following figure. The log odds are displayed as –5 for clinics 5 and 6 at the 0 treatment level, although the actual log odds are –∞. The log odds increase from the 0 to 1 level of the treatment in all but one clinic (clinic 8), so we would expect the estimate of β_1 to be positive.



The PROC NLMIXED statements to fit the logistic-normal model to these data are as follows:

```
proc nlmixed data=infection;
   parms beta0=-1 beta1=1 s2u=2;
   eta = beta0 + beta1*t + u;
   expeta = exp(eta);
   p = expeta/(1+expeta);
   model x ~ binomial(n,p);
   random u ~ normal(0,s2u)
      subject=clinic;
   predict eta out=eta;
   estimate '1/beta1' 1/beta1;
run;
```

The PROC NLMIXED statement invokes the procedure, and the PARMS statement defines the parameters and their starting values. Reasonable starting values such as these can often be obtained by fitting a simpler model.

The next three statements construct the variable P to correspond to the p_{ij}, and the MODEL statement defines the conditional distribution of x_i to be binomial. The RANDOM statement defines U to be the random effect with subjects defined by the CLINIC variable.

The PREDICT statement constructs predictions for each observation in the input data set. For this example, predictions of n_i are output to a SAS data set named ETA. These predictions are linear functions of the empirical Bayes estimates of the random effects u_i. The ESTIMATE statement requests an estimate of the reciprocal of β_1.

The output for this model is as follows.

```
                The NLMIXED Procedure

                   Specifications

Data Set                              WORK.
                                      INFECTION
Dependent Variable                    x
Distribution for Dep Variable         Binomial
Random Effects                        u
Distribution for Random Effects       Normal
Subject Variable                      clinic
Optimization Technique                Dual Quasi-
                                      Newton
Estimation Method                     Adaptive
                                      Gaussian
                                      Quadrature
```

The “Specifications” table provides basic information about the nonlinear mixed model.

```
                     Dimensions

   Observations Used                    16
   Observations Not Used                 0
   Total Observations                   16
   Subjects                              8
   Max Obs Per Subject                   2
   Parameters                            3
   Quadrature Points                     5
```

The “Dimensions” table provides counts of various variables. You should check this table to make sure the data set and model have been entered properly. PROC NLMIXED selects five quadrature points to achieve the default accuracy in the likelihood calculations.

Parameters			
beta0	beta1	s2u	NegLogLike
-1	1	2	37.5945925

The "Parameters" table lists the starting point of the optimization.

Iterations

Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	2	37.3622692	0.232323	2.882077	-19.3762
2	3	37.1460375	0.216232	0.921926	-0.82852
3	5	37.0300936	0.115944	0.315897	-0.59175
4	6	37.0223017	0.007792	0.01906	-0.01615
5	7	37.0222472	0.000054	0.001743	-0.00011
6	9	37.0222466	6.57E-7	0.000091	-1.28E-6
7	11	37.0222466	5.38E-10	2.078E-6	-1.1E-9

NOTE: GCONV convergence criterion satisfied.

The "Iterations" table indicates successful convergence in seven iterations.

Fitting Information

-2 Log Likelihood	74.0
AIC (smaller is better)	80.0
BIC (smaller is better)	80.3
Log Likelihood	-37.0
AIC (larger is better)	-40.0
BIC (larger is better)	-40.1

The "Fitting Information" table lists some useful statistics based on the maximized value of the log likelihood.

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|
beta0	-1.1974	0.5561	7	-2.15	0.0683
beta1	0.7385	0.3004	7	2.46	0.0436
s2u	1.9591	1.1903	7	1.65	0.1438

Parameter	Alpha	Lower	Upper	Gradient
beta0	0.05	-2.5123	0.1175	-3.1E-7
beta1	0.05	0.02806	1.4488	-2.08E-6
s2u	0.05	-0.8554	4.7736	-2.48E-7

The "Parameter Estimates" table indicates marginal significance of the two fixed-effects parameters. The positive value of the estimate of β_1 indicates that the treatment significantly increases the chance of a favorable cure.

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
1/beta1	1.3542	0.5509	7	2.46	0.0436

Alpha	Lower	Upper
0.05	0.05146	2.6569

The "Additional Estimates" table displays results from the ESTIMATE statement. The estimate of $1/\beta_1$ equals 1/0.7385 = 1.3542 and its standard error equals $0.3004/0.7385^2$= 0.5509 by the delta method (Billingsley 1986). Note this particular approximation produces a *t* statistic that is identical to that for the estimate of β_1 itself.

Not shown is the ETA data set, which contains the original 16 observations, predictions of the n_{ij}, and associated statistics.

SYNTAX

This section provides an overview of the statements that are available in PROC NLMIXED and some of their key options. More details and additional options are provided in the complete documentation referenced in the CONTACT INFORMATION section at the end of this paper.

PROC NLMIXED *options*;

This statement invokes the procedure. A large number of options are available; the following are some of the most important ones:

ALPHA= specifies the alpha level used to compute t statistics and intervals.

COV requests the approximate covariance matrix for the parameter estimates.

CORR requests the approximate correlation matrix of the parameter estimates.

DATA= specifies the input data set.

ECOV requests the approximate covariance matrix for all quantities specified in ESTIMATE statements.

ECORR requests the approximate covariance matrix for all quantities specified in ESTIMATE statements.

FD requests finite difference derivatives.

GCONV= specifies the relative gradient convergence criterion.

HESS requests the display of the final Hessian matrix.

ITDETAILS requests a more detailed iteration history.

MAXITER= specifies the maximum number of iterations.

METHOD= specifies the method for approximating the integral over the random effects. Valid values are FIRO, GAUSS (the default), HARDY, and ISAMP.

QPOINTS= specifies the number of quadrature points for each random effect.

START requests the printout of the starting gradient.

TECH= specifies the optimization technique. Valid values are CONGRA, DBLDOG, NMSIMP, NONE, NEWRAP, NRRIDG, QUANEW (the default), and TRUREG.

ARRAY ***arrayname***;

The ARRAY statement allows you to specify SAS arrays.

BOUNDS ***b_con [, b_con];***

The BOUNDS statement enables you to specify boundary constraints on the parameters. Example statements are as follows:

```
bounds 0 <= a1-a9 X <= 1, -1 <= c2-c5;
bounds b1-b10 y >= 0;
```

BY ***variables***;

The BY statement invokes standard SAS BY processing.

ESTIMATE ***'label' expression*****;**

The ESTIMATE statement enables you to compute an additional estimate that is a function of the parameter values. Multiple ESTIMATE statements are allowed.

ID ***expressions***;

The ID statement identifies additional quantities to be included in the OUT= data set of all PREDICT statements.

MODEL ***dependent-variable ~ distribution***;

The MODEL statement specifies the dependent variable and its conditional distribution given the random effects. Valid distributional forms are NORMAL(m,v), BINARY(p), BINOMIAL(n,p), POISSON(m), and GENERAL(ll).

PARMS ***<name_list[=numbers]>;***

The PARMS statement specifies parameter names and their starting values.

PREDICT ***expression OUT=SAS-data-set***;

The PREDICT statement enables you to construct predictions for every observation in the input data set. Multiple PREDICT statements are allowed.

RANDOM ***random-effects ~ distribution SUBJECT=variable options;***

The RANDOM statement specifies the random effects and their distribution. The only currently available distribution is NORMAL. The SUBJECT= argument specifies a variable indicating subjects. One option is OUT=, giving a SAS data set name to contain empirical Bayes estimates of the random effects. Example statements are as follows:

```
random u ~ normal(0,s2u) subject=clinic;
random b1 b2 ~ normal([0,0],[s2b1,cb12,s2b2])
      subject=person out=eb;
```

Program Statements

Standard SAS programming statements enable you to construct a wide variety of nonlinear models. Besides usual assignment and mathematical function statements, you may also use CALL, DO, GOTO, IF, PUT, and WHEN statements.

OUTPUT TABLE NAMES

PROC NLMIXED uses the Version 7 Output Delivery System. To convert any table to a SAS data set, use the ODS statement in the following form:

```
ods output 'table-name'=SAS-data-set;
```

Here, TABLE-NAME is one of the names from the table below, and SAS-DATA-SET is a name you select for the output SAS data set.

Table Name	Statement / Option
AdditionalEstimates	ESTIMATE
ConvergenceStatus	default
CorrMatAddEst	ECORR
CorrMatParmEst	CORR
CovMatAddEst	ECOV
CovMatParmEst	COV
Dimensions	default
Fitting	default
Hessian	HESS
Iterations	default
Parameters	default
ParameterEstimates	default
Specifications	default
StartingHessian	START HESS
StartingValues	START

To suppress all displayed output, use

```
ods exclude all;
```

and use

```
ods select all;
```

to redisplay it.

LIMITATIONS

PROC NLMIXED is best suited for models with a single random effect, although you can also successfully compute integrals in two and three dimensions as well. In addition, you can use PROC NLMIXED as a general optimization tool by omitting random effects altogether. Problems which are badly scaled or sufficiently noisy will not perform well with PROC NLMIXED. Also, PROC NLMIXED currently does not generally handle nested or crossed random effects.

CONCLUSION

PROC NLMIXED offers a flexible and powerful environment for fitting nonlinear mixed models. Its main computing components are a SAS engine for processing programming statements, a numerical integrator, and a library of optimization routines. These tools enable you to fit many common nonlinear and generalized linear mixed models using likelihood-based methods.

REFERENCES

Anderson, D.A. and Aitkin, M. (1985), "Variance Component Models with Binary Response: Interviewer Variability," *Journal of the Royal Statistical Society B*, 47, 203-210.

Beal, S.L. and Sheiner, L.B. (1982), "Estimating Population Kinetics," *CRC Crit. Rev. Biomed. Eng.*, 8, 195-222.

Beal, S.L. and Sheiner, L.B. (1988), "Heteroskedastic Nonlinear Regression," *Technometrics*, 30, 327-338.

Beal, S.L. and Sheiner, L.B., eds. (1992), *NONMEM User's Guide*, University of California, San Francisco, NONMEM Project Group.

Beitler, P.J. and Landis, J.R. (1985), "A Mixed-effects Model for Categorical Data," *Biometrics*, 41, 991-1000.

Billingsley, P. (1986), *Probability and Measure, Second Edition,* New York: John Wiley & Sons, Inc.

Booth, J.G. and Hobert, J.P. (1998), "Standard Errors of Prediction in Generalized Linear Mixed Models," *Journal of the American Statistical Association*, 93, 262-272.

Breslow, N.E. and Clayton, D.G. (1993), "Approximate Inference in Generalized Linear Mixed Models," *Journal of the American Statistical Association*, 88, 9-25.

Crouch, E.A.C. and Spiegelman, D. (1990), "The Evaluation of Integrals of the Form $\int_{-\infty}^{\infty} f(t)\exp(-t^2)dt$: Application to Logistic-normal Models," *Journal of the American Statistical Association*, 85, 464-469

Davidian, M. and Gallant, R.A. (1993), "The Nonlinear Mixed Effects Model with a Smooth Random Effects Density," *Biometrika*, 80, 475-488.

Davidian, M. and Giltinan, D.M. (1995*), Nonlinear Models for Repeated Measurement Data*, New York: Chapman & Hall.

Diggle, P.J., Liang, K.Y., and Zeger, S.L. (1994), *Analysis of Longitudinal Data*, Oxford: Clarendon Press.

Draper, N.R. and Smith, H. (1981), *Applied Regression Analysis, Second Edition*, New York: John Wiley & Sons, Inc.

Engel, B. and Keen, A. (1992), "A Simple Approach for the Analysis of Generalized Linear Mixed Models," LWA-92-6, Agricultural Mathematics Group (GLW-DLO). Wageningen, The Netherlands.

Gilmour, A.R., Anderson, R.D., and Rae, A.L. (1985), "The Analysis of Binomial Data by Generalized Linear Mixed Model," *Biometrika*, 72, 593-599.

Goldstein, H. (1991), "Nonlinear Multilevel Models, with an Application to Discrete Response Data," *Biometrika*, 78, 45-51.

Harville, D.A. and Mee, R.W. (1984), "A Mixed-model Procedure for Analyzing Ordered Categorical Data," *Biometrics*, 40, 393-408.

Hedeker, D. and Gibbons, R.D. (1994), "A Random Effects Ordinal Regression Model for Multilevel Analysis," *Biometrics*, 50, 933-944.

Lin, X. and Breslow, N.E. (1996), "Bias Correction in Generalized Linear Mixed Models with Multiple Components of Dispersion," *Journal of the American Statistical Association*, 91, 1007-1016.

Lindstrom, M.J. and Bates, D.M. (1990), "Nonlinear Mixed Effects Models for Repeated Measures Data," *Biometrics*, 46, 673-687.

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. (1996), *SAS System for Mixed Model*s, Cary, NC: SAS Institute Inc.

Liu, Q. and Pierce, D.A. (1994), "A Note on Gauss-Hermite Quadrature," *Biometrika*, 81, 624-629.

Longford, N.T. (1994), "Logistic Regression with Random Coefficients," *Computational Statistics and Data Analysis*, 17, 1-15.

McCulloch, C.E. (1994), "Maximum Likelihood Variance Components Estimation for Binary Data," *Journal of the American Statistical Association*, 89, 330-335.

McGilchrist, C.E. (1994), "Estimation in Generalized Mixed Models," *Journal of the Royal Statistical Society B*, 56, 61-69.

Ochi, Y. and Prentice, R.L. (1984), "Likelihood Inference in a Correlated Probit Regression Model," *Biometrika*, 71, 531-543.

Pierce, D.A. and Sands, B.R. (1975), "Extra-Bernoulli Variation in Binary Data," Technical Report 46, Department of Statistics, Oregon State University.

Pinheiro, J.C. and Bates, D.M. (1995), "Approximations to the Log-likelihood Function in the Nonlinear Mixed-effects Model," *Journal of Computational and Graphical Statistics*, 4, 12-35.

Rodriguez, G. and Goldman, N. (1995), "An Assessment of Estimation Procedures for Multilevel Models with Binary Response," *Journal of the Royal Statistical Society A*, 158, 73-89.

Schall, R. (1991). "Estimation in Generalized Linear Models with Random Effects," *Biometrika*, 78, 719-727.

Sheiner L. B. and Beal S. L., "Evaluation of Methods for Estimating Population Pharmacokinetic Parameters. I. Michaelis-Menten Model: Routine Clinical Pharmacokinetic Data," *Journal of Pharmacokinetics and Biopharmaceutics*, 8, (1980) 553-571.

Sheiner, L.B. and Beal, S.L. (1985), "Pharmacokinetic Parameter Estimates from Several Least Squares Procedures: Superiority of Extended Least Squares," *Journal of Pharmacokinetics and Biopharmaceutics*, 13, 185-201.

Stiratelli, R., Laird, N.M., and Ware, J.H. (1984), "Random Effects Models for Serial Observations with Binary Response," *Biometrics*, 40, 961-971.

Vonesh, E.F., (1992), "Nonlinear Models for the Analysis of Longitudinal Data," *Statistics in Medicine*, 11, 1929-1954.

Vonesh, E.F., (1996), "A Note on Laplace's Approximation in Nonlinear Mixed Effects Models," *Biometrika*, 83, 447-452.

Vonesh, E.F. and Chinchilli, V.M. (1996), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, NewYork: Marcel Dekker.

Wolfinger R.D. (1993), "Laplace's Approximation for Nonlinear Mixed Models," *Biometrika*, 80, 791-795.

Wolfinger, R.D. (1997), "Comment: Experiences with the SAS Macro NLINMIX," *Statistics in Medicine*, 16, 1258-1259.

Wolfinger, R.D., and Lin, X. (1997), "Two Taylor-series Approximation Methods for Nonlinear Mixed Models," *Computational Statistics and Data Analysis*, 25, 465-490.

Wolfinger, R.D. and O'Connell, M. (1993), "Generalized Linear Mixed Models: a Pseudo-likelihood Approach," *Journal of Statistical Computation and Simulation*, 48, 233-243.

CONTACT INFORMATION

PROC NLMIXED requires Version 7 of the SAS System. Complete documentation is available on the web at *http://www.sas.com/techsup/download/stat/* in the Postscript file *nlmixed.ps*. You may send feedback to me at *sasrdw@sas.com*.