Directions: All students are to answer both exercises below showing all relevant work; conclusions and justifications are to be given in clear detailed English. Please type up your solutions or write very neatly.

1. Diggle, Liang & Zeger (1994) and Venables and Ripley 1999, p.209 give an example of repeated measurements on the '**size**' – defined as log(height) + 2 log(diameter) – of 79 Sitka spruce trees, 54 of which were grown in ozone-enriched chambers and 25 of which were controls. The trees were measured on eight different days in 1989 (see the SAS analysis in the attached *Appendix*). This analysis uses the classical linear mixed models analysis. Use the output to thoroughly analyze these repeated-measures data.
   (a) For each part of the analysis (i.e., for each of the PROC MIXEDs), listing the necessary assumptions paying particular attention to the assumed form of the variance-covariance matrix.
   (b) Identify and justify which one of the three covariance structures used in the PROC MIXEDs is most appropriate for these data. 'Justify' here means report the relevant test statistic(s) and results. Clearly specify how many variance components are estimated for each of the three runs of PROC MIXED.
   (c) Using the appropriate analysis, comment on whether you feel the profiles for the two treatment curves are the same (some people call the curves "parallel"), reporting the relevant test statistic and p-value. Give your final conclusion.

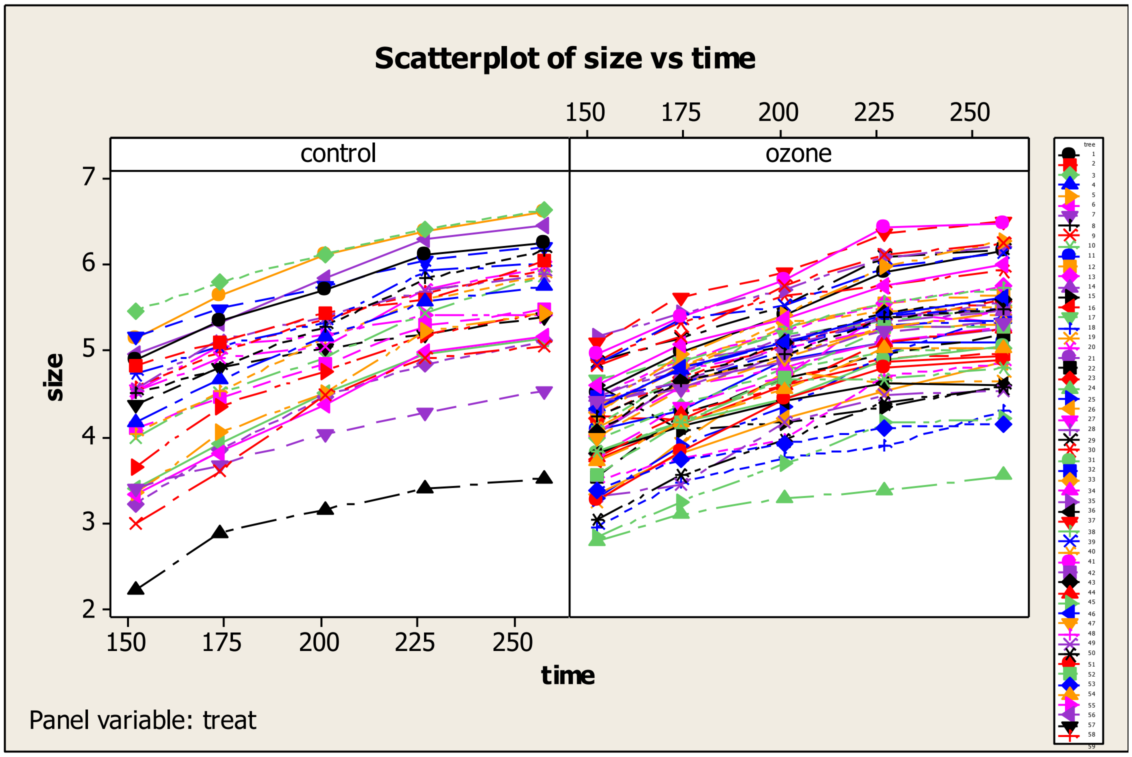
2. On page 18 of Davidian & Giltinan, *Nonlinear Models for Repeated Measurement Data*, the authors present the data of Kwan et al (1976). Here, we fit the mixed nonlinear model of the form

$$E(Y) = \beta_1 e^{-\beta_2 x} + \beta_3 e^{-\beta_4 x}$$

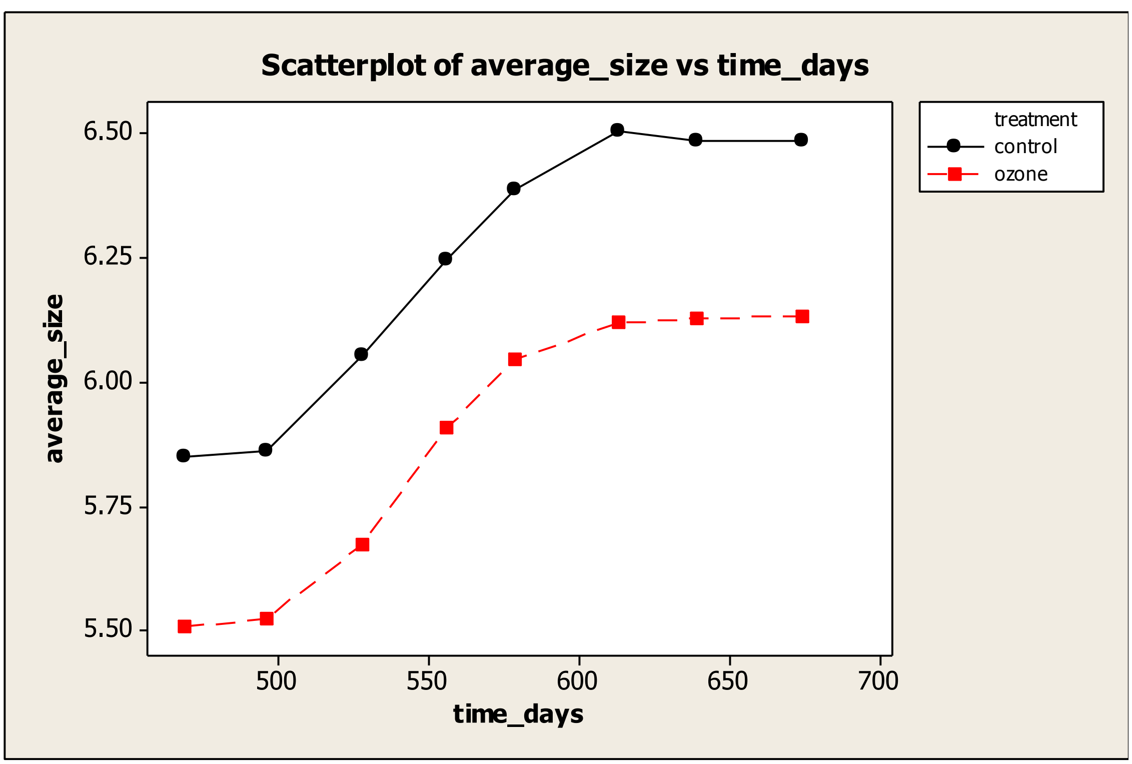
   In this expression, Y = plasma concentration of the drug 'cefamandole' and x = time (in minutes post dosing). For this study, a dose of 15 mg (per kg body weight) of the drug was administered by ten-minute intravenous infusion to six healthy volunteers. The data are analyzed in SAS in the *Appendix.* The program first graphs the data then runs four runs of PROC NLMIXED. To understand the model function, the above four-parameter exponential (SE4) model function is graphed in the Appendix (p.9) using the values of β_1 = 2.7733, β_2 = 2.8139, β_3 = 0.7870, β_4 = 0.4195; also graphed are the two components, $[\beta_1 e^{-\beta_2 x}]$ and $[\beta_3 e^{-\beta_4 x}]$. For this model function, note that the 'upper asymptote' is $\beta_1 + \beta_3$; also β_2 is the initial slope parameter; and β_4 is the final slope parameter.
   (a) Listing all necessary assumptions, explain what is being done in each of the four runs of NLMIXED, commenting on which model is being fit and identifying the underlying assumptions. Be specific, and in each case write out the assumed covariance matrix for the model parameters. Also, identify which models are special cases of others (i.e., which are nested), and identify which they are nested in.
   (b) Choose the NLMIXED model/analysis which **best** describes these data, and give specific reasons for why you chose the model you did, and why you rejected each of the others. Give test statistics, degrees of freedom and p-values to justify your claims.
   (c) The parameters β_2 and β_4 are important since they address the rate of decrease (elimination) of the expected concentration function. Contrasting the estimated standard errors (SE's) of these two parameters for the first NLMIXED with those for each of the second through fourth NLMIXED's, why are these SEs lower for the latter three NLMIXED's than for the first NLMIXED?

Homework 13 Appendix

Graphs for First Exercise



Graphs of means for both treatments (average size versus time)



First Exercise – MIXED (CS) Approach

```
proc mixed data=sitka89;
  title 'MIXED approach with CS error structure';
  class treat tree time;
  model size=treat time treat*time;
  repeated / type=cs sub=tree(treat);
run;
```

```
                    MIXED approach with CS error structure

                            The Mixed Procedure
                             Model Information

          Data Set                      WORK.SITKA89
          Dependent Variable            size
          Covariance Structure          Compound Symmetry
          Subject Effect                tree(treat)
          Estimation Method             REML
          Residual Variance Method      Profile
          Fixed Effects SE Method       Model-Based
          Degrees of Freedom Method     Between-Within

                         Class Level Information
            Class    Levels    Values
            treat         2    control ozone
            tree         79    1 2 3 4 5 6 7 8 9 10 11 12 13
                               14 15 16 17 18 19 20 21 22 23
                               24 25 26 27 28 29 30 31 32 33
                               34 35 36 37 38 39 40 41 42 43
                               44 45 46 47 48 49 50 51 52 53
                               54 55 56 57 58 59 60 61 62 63
                               64 65 66 67 68 69 70 71 72 73
                               74 75 76 77 78 79
            time          8    469 496 528 556 579 613 639
                               674

                               Dimensions
                 Covariance Parameters            2
                 Columns in X                    27
                 Columns in Z                     0
                 Subjects                        79
                 Max Obs Per Subject              8

                         Number of Observations
               Number of Observations Read             632
               Number of Observations Used             632
               Number of Observations Not Used           0

                            The Mixed Procedure
                             Iteration History
       Iteration    Evaluations    -2 Res Log Like       Criterion
               0              1      1260.72452373
               1              1      -614.89217445      0.00000000
                        Convergence criteria met.

                      Covariance Parameter Estimates
```

```
                   Cov Parm     Subject        Estimate
                   CS           tree(treat)      0.4033
                   Residual                    0.009478

                             Fit Statistics
                  -2 Res Log Likelihood          -614.9
                  AIC (smaller is better)        -610.9
                  AICC (smaller is better)       -610.9
                  BIC (smaller is better)        -606.2

                     Null Model Likelihood Ratio Test
                       DF    Chi-Square      Pr > ChiSq
                        1       1875.62          <.0001

                      Type 3 Tests of Fixed Effects
                            Num     Den
             Effect          DF      DF    F Value    Pr > F
             treat            1      77       5.30    0.0241
             time             7     539     545.99    <.0001
             treat*time       7     539       0.66    0.7071
```

First Exercise – MIXED (AR1) Approach

```
proc mixed data=sitka89;
  title 'MIXED approach with AR(1) error structure';
  class treat tree time;
  model size=treat time treat*time;
  repeated / type=ar(1) sub=tree(treat);
run;
```

```
           MIXED approach with AR(1) error structure

                         The Mixed Procedure
                          Model Information
           Data Set                     WORK.SITKA89
           Dependent Variable           size
           Covariance Structure         Autoregressive
           Subject Effect               tree(treat)
           Estimation Method            REML
           Residual Variance Method     Profile
           Fixed Effects SE Method      Model-Based
           Degrees of Freedom Method    Between-Within

                        Class Level Information
             Class    Levels    Values
             treat         2    control ozone
             tree         79    1 2 3 4 5 6 7 8 9 10 11 12 13
                                14 15 16 17 18 19 20 21 22 23
                                24 25 26 27 28 29 30 31 32 33
                                34 35 36 37 38 39 40 41 42 43
                                44 45 46 47 48 49 50 51 52 53
                                54 55 56 57 58 59 60 61 62 63
                                64 65 66 67 68 69 70 71 72 73
                                74 75 76 77 78 79
             time          8    469 496 528 556 579 613 639
                                674
```

```
                         Dimensions
            Covariance Parameters            2
            Columns in X                    27
            Columns in Z                     0
            Subjects                        79
            Max Obs Per Subject              8

                   Number of Observations
        Number of Observations Read             632
        Number of Observations Used             632
        Number of Observations Not Used           0

                      Iteration History
Iteration    Evaluations    -2 Res Log Like       Criterion
        0              1      1260.72452373
        1              2      -892.46138132      0.00114345
        2              1      -893.74611704      0.00004058
        3              1      -893.78818880      0.00000005
        4              1      -893.78824278      0.00000000
                  Convergence criteria met.

               Covariance Parameter Estimates
            Cov Parm     Subject        Estimate
            AR(1)        tree(treat)      0.9911
            Residual                      0.4272

                       Fit Statistics
           -2 Res Log Likelihood          -893.8
           AIC (smaller is better)        -889.8
           AICC (smaller is better)       -889.8
           BIC (smaller is better)        -885.0

              Null Model Likelihood Ratio Test
                DF    Chi-Square      Pr > ChiSq
                 1       2154.51          <.0001

               Type 3 Tests of Fixed Effects
                      Num     Den
      Effect           DF      DF    F Value    Pr > F
      treat             1      77       5.13    0.0263
      time              7     539     135.39    <.0001
      treat*time        7     539       2.03    0.0500
```

First Exercise – MIXED (UN) Approach

```
proc mixed data=sitka89;
  title 'MIXED approach with Unstructured error structure';
  class treat tree time;
  model size=treat time treat*time;
  repeated / type=un sub=tree(treat);
run;
```

```
          MIXED approach with Unstructured error structure

                        The Mixed Procedure
```

Model Information

Data Set	WORK.SITKA89
Dependent Variable	size
Covariance Structure	Unstructured
Subject Effect	tree(treat)
Estimation Method	REML
Residual Variance Method	None
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
treat	2	control ozone
tree	79	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79
time	8	469 496 528 556 579 613 639 674

Dimensions

Covariance Parameters	36
Columns in X	27
Columns in Z	0
Subjects	79
Max Obs Per Subject	8

Number of Observations

Number of Observations Read	632
Number of Observations Used	632
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	1260.72452373	
1	1	-1246.66153301	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

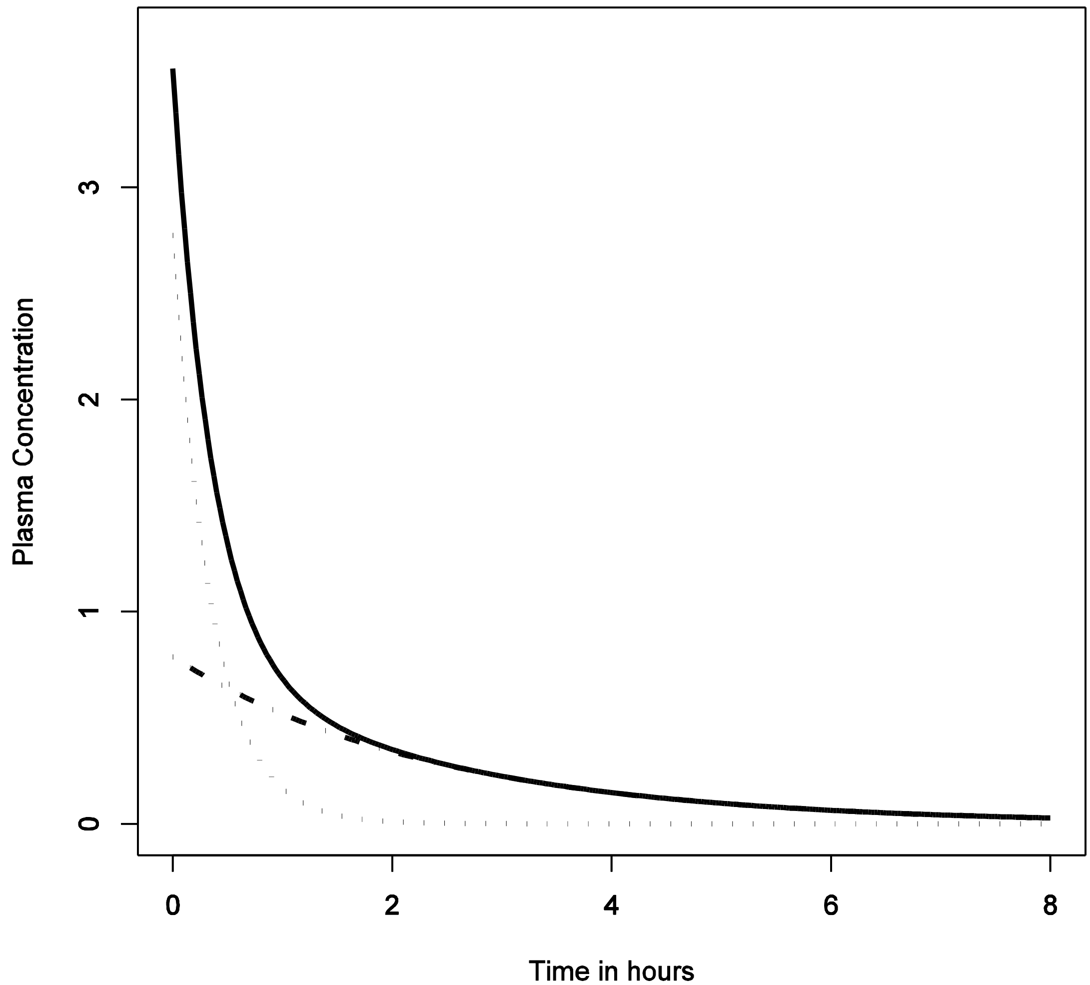
Cov Parm	Subject	Estimate
UN(1,1)	tree(treat)	0.4486
UN(2,1)	tree(treat)	0.4455
UN(2,2)	tree(treat)	0.4425
UN(3,1)	tree(treat)	0.4189
UN(3,2)	tree(treat)	0.4166
UN(3,3)	tree(treat)	0.4012
UN(4,1)	tree(treat)	0.4080
UN(4,2)	tree(treat)	0.4060
UN(4,3)	tree(treat)	0.3890
UN(4,4)	tree(treat)	0.3877
UN(5,1)	tree(treat)	0.4223
UN(5,2)	tree(treat)	0.4201
UN(5,3)	tree(treat)	0.4019
UN(5,4)	tree(treat)	0.4004
UN(5,5)	tree(treat)	0.4229
UN(6,1)	tree(treat)	0.4119

```
UN(6,2)     tree(treat)     0.4099
UN(6,3)     tree(treat)     0.3944
UN(6,4)     tree(treat)     0.3930
UN(6,5)     tree(treat)     0.4115
UN(6,6)     tree(treat)     0.4061
UN(7,1)     tree(treat)     0.3986
UN(7,2)     tree(treat)     0.3967
UN(7,3)     tree(treat)     0.3807
UN(7,4)     tree(treat)     0.3790
UN(7,5)     tree(treat)     0.3950
UN(7,6)     tree(treat)     0.3909
UN(7,7)     tree(treat)     0.3851
UN(8,1)     tree(treat)     0.4085
UN(8,2)     tree(treat)     0.4065
UN(8,3)     tree(treat)     0.3918
UN(8,4)     tree(treat)     0.3909
UN(8,5)     tree(treat)     0.4077
UN(8,6)     tree(treat)     0.4030
UN(8,7)     tree(treat)     0.3936
UN(8,8)     tree(treat)     0.4079

           Fit Statistics
-2 Res Log Likelihood        -1246.7
AIC (smaller is better)      -1174.7
AICC (smaller is better)     -1170.1
BIC (smaller is better)      -1089.4

   Null Model Likelihood Ratio Test
     DF    Chi-Square      Pr > ChiSq
     35       2507.39          <.0001

     Type 3 Tests of Fixed Effects
                 Num     Den
Effect            DF      DF    F Value    Pr > F
treat              1      77       5.30    0.0241
time               7      77     141.48    <.0001
treat*time         7      77       2.15    0.0485
```

<u>The fitted variance-covariance for these data follows</u>

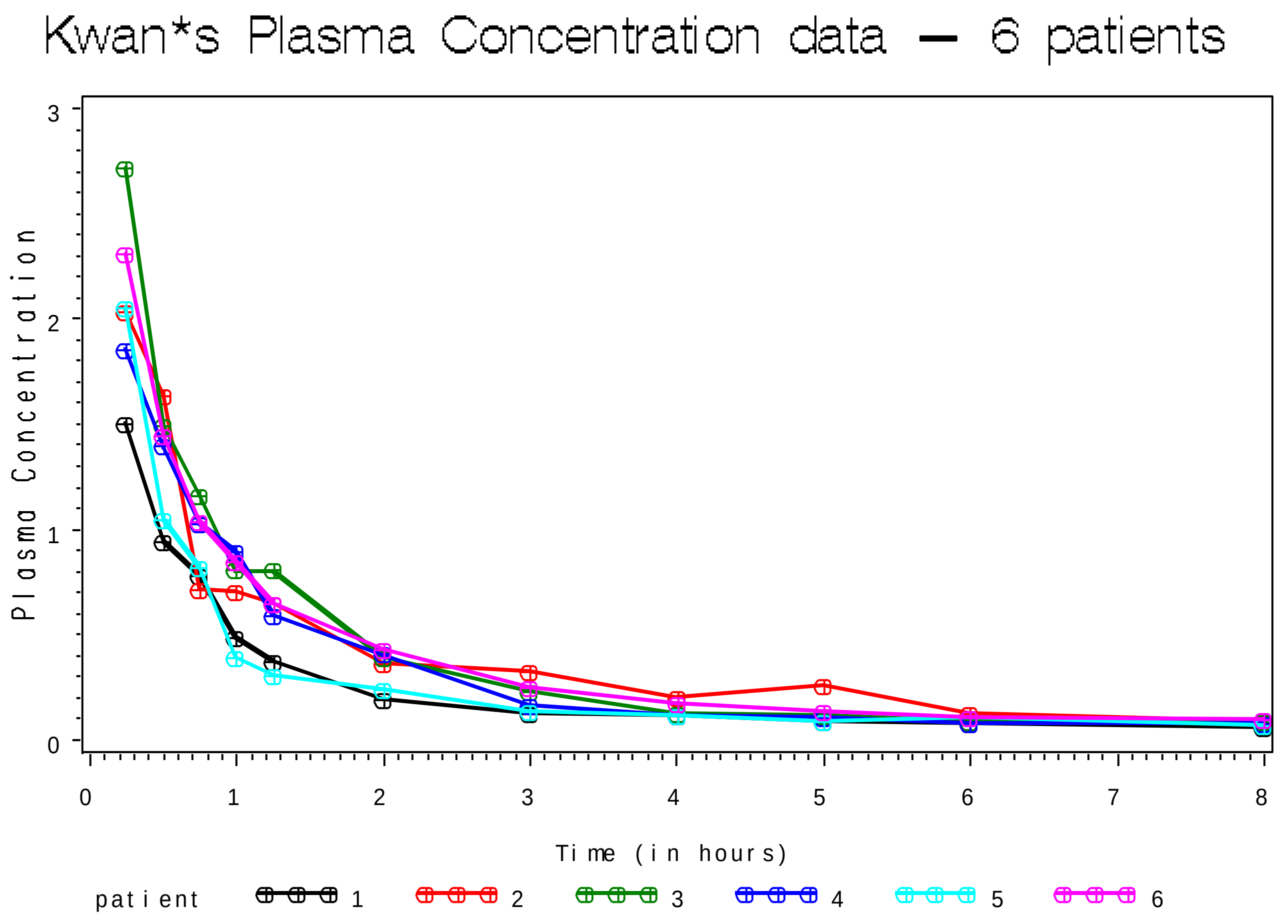
0.4486							
0.4455	0.4425	*This matrix is symmetric*					
0.4189	0.4166	0.4012					
0.4080	0.4060	0.3890	0.3877				
0.4223	0.4201	0.4019	0.4004	0.4229			
0.4119	0.4099	0.3944	0.3930	0.4115	0.4061		
0.3986	0.3967	0.3807	0.3790	0.3950	0.3909	0.3851	
0.4085	0.4065	0.3918	0.3909	0.4077	0.4030	0.3936	0.4079

Second Exercise Plot of Theoretical Model Function (upper-most curve)



Second Exercise Plot of Data for 6 Patients

```
data one;
  do patient=1 to 6;
  do time=.25 to 1.25 by .25,2,3,4,5,6,8;
    input y @@; plasma=y/100; drop y; output;
  end; end; cards;
150  94  78 48 37 19 12 11  8  7 5 203 163  71 70 64 36 32 20 25 12 8
272 149 116 80 80 39 22 12 11  8 8 185 139 102 89 59 40 16 11 10  7 7
205 104  81 39 30 23 13 11  8 10 6 231 144 103 84 64 42 24 17 13 10 9
;
axis1 label=(a=90 h=1.2 'Plasma Concentration');
axis2 label=(h=1.2 'Time (in hours)');
symbol1 v=+ i=join;   symbol2 v=x i=join;      symbol3 v=* i=join;
symbol4 v=dot i=join; symbol5 v=circle i=join; symbol6 v=$ i=join;
proc gplot;
  title 'Kwan*s Plasma Concentration data - 6 patients';
  plot plasma*time=patient / vaxis=axis1 haxis=axis2;
run;
```



Second Exercise – First NLMIXED

```
proc nlmixed;
  parms b1=3.5 b2=3.6 b3=.5 b4=.5 s2e=.1;
  rhs=b1*exp(-b2*time)+b3*exp(-b4*time);
  model plasma~normal(rhs,s2e);
run;
```

```
                         The NLMIXED Procedure
                            Specifications
Data Set                                    WORK.ONE
Dependent Variable                          plasma
Distribution for Dependent Variable         Normal
Optimization Technique                      Dual Quasi-Newton
Integration Method                          None

                              Dimensions
            Observations Used                     66
            Observations Not Used                  0
            Total Observations                    66
            Parameters                             5
```

```
                          Fit Statistics
              -2 Log Likelihood                 -47.3
              AIC (smaller is better)           -37.3
              AICC (smaller is better)          -36.3
              BIC (smaller is better)           -26.3

                        Parameter Estimates
                     Standard
Parameter  Estimate     Error    DF  t Value  Pr > |t|   Alpha     Lower     Upper
b1           2.7734    0.2485    66    11.16    <.0001    0.05    2.2773    3.2695
b2           2.4263    0.5997    66     4.05    0.0001    0.05    1.2289    3.6236
b3           0.6067    0.3010    66     2.02    0.0479    0.05  0.005707    1.2078
b4           0.3356    0.1542    66     2.18    0.0331    0.05   0.02778    0.6434
s2e         0.02860  0.004979    66     5.74    <.0001    0.05   0.01866   0.03854
```

Second Exercise – Second NLMIXED

```
proc nlmixed;
  parms b1=3.5 b2=3.6 b3=.5 b4=.5 s2e=.1 s2b1=.1 s2b3=.1 cb13=0;
  rhs=(b1+u1)*exp(-b2*time)+(b3+u3)*exp(-b4*time);
  model plasma~normal(rhs,s2e);
  random u1 u3~normal([0,0],[s2b1,cb13,s2b3]) subject=patient;
run;
```

```
                        The NLMIXED Procedure
                           Specifications
          Data Set                                    WORK.ONE
          Dependent Variable                          plasma
          Distribution for Dependent Variable         Normal
          Random Effects                              u1 u3
          Distribution for Random Effects             Normal
          Subject Variable                            patient
          Optimization Technique                      Dual Quasi-Newton
          Integration Method                          Adaptive Gaussian
                                                      Quadrature

                             Dimensions
              Observations Used                    66
              Observations Not Used                 0
              Total Observations                   66
              Subjects                              6
              Max Obs Per Subject                  11
              Parameters                            8
              Quadrature Points                     1

                           Fit Statistics
              -2 Log Likelihood                -112.2
              AIC (smaller is better)           -96.2
              AICC (smaller is better)          -93.7
              BIC (smaller is better)           -97.9

                        Parameter Estimates
                     Standard
```

Parameter	Estimate	Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
b1	2.7733	0.2761	4	10.04	0.0006	0.05	2.0066	3.5400
b2	2.8139	0.3600	4	7.82	0.0014	0.05	1.8144	3.8135
b3	0.7870	0.1731	4	4.55	0.0104	0.05	0.3064	1.2676
b4	0.4195	0.07116	4	5.90	0.0041	0.05	0.2220	0.6171
s2e	0.006739	0.001298	4	5.19	0.0066	0.05	0.003134	0.01034
s2b1	0.3346	0.2338	4	1.43	0.2256	0.05	-0.3145	0.9837
s2b3	0.04391	0.03076	4	1.43	0.2266	0.05	-0.04148	0.1293
cb13	-0.00179	0.05782	4	-0.03	0.9767	0.05	-0.1623	0.1587

Second Exercise – Third NLMIXED

```
proc nlmixed;
  parms b1=3.5 b2=3.6 b3=.5 b4=.5 s2e=.1 s2b1=.1;
  rhs=(b1+u1)*exp(-b2*time)+b3*exp(-b4*time);
  model plasma~normal(rhs,s2e);
  random u1~normal(0,s2b1) subject=patient;
run;
```

The NLMIXED Procedure

Specifications

Data Set	WORK.ONE
Dependent Variable	plasma
Distribution for Dependent Variable	Normal
Random Effects	u1
Distribution for Random Effects	Normal
Subject Variable	patient
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions

Observations Used	66
Observations Not Used	0
Total Observations	66
Subjects	6
Max Obs Per Subject	11
Parameters	6
Quadrature Points	1

Fit Statistics

-2 Log Likelihood	-92.9
AIC (smaller is better)	-80.9
AICC (smaller is better)	-79.4
BIC (smaller is better)	-82.1

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
b1	2.7617	0.2816	5	9.81	0.0002	0.05	2.0377	3.4856
b2	2.1108	0.2473	5	8.53	0.0004	0.05	1.4750	2.7467
b3	0.4522	0.1426	5	3.17	0.0248	0.05	0.08557	0.8189
b4	0.2617	0.08520	5	3.07	0.0277	0.05	0.04271	0.4807
s2e	0.01098	0.002005	5	5.48	0.0028	0.05	0.005828	0.01614

s2b1	0.3680	0.2330	5	1.58	0.1750	0.05	-0.2309	0.9669

Second Exercise – Fourth NLMIXED

```
proc nlmixed;
  parms b1=3.5 b2=3.6 b3=.5 b4=.5 s2e=.1 s2b1=.1 s2b3=.1;
  rhs=(b1+u1)*exp(-b2*time)+(b3+u3)*exp(-b4*time);
  model plasma~normal(rhs,s2e);
  random u1 u3~normal([0,0],[s2b1,0,s2b3]) subject=patient;
run;
```

The NLMIXED Procedure

Specifications

Data Set	WORK.ONE
Dependent Variable	plasma
Distribution for Dependent Variable	Normal
Random Effects	u1 u3
Distribution for Random Effects	Normal
Subject Variable	patient
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

Dimensions

Observations Used	66
Observations Not Used	0
Total Observations	66
Subjects	6
Max Obs Per Subject	11
Parameters	7
Quadrature Points	1

Fit Statistics

-2 Log Likelihood	-112.2
AIC (smaller is better)	-98.2
AICC (smaller is better)	-96.3
BIC (smaller is better)	-99.7

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
b1	2.7736	0.2756	4	10.06	0.0005	0.05	2.0085	3.5388
b2	2.8132	0.3594	4	7.83	0.0014	0.05	1.8155	3.8110
b3	0.7865	0.1724	4	4.56	0.0103	0.05	0.3080	1.2651
b4	0.4192	0.07058	4	5.94	0.0040	0.05	0.2233	0.6152
s2e	0.006741	0.001297	4	5.20	0.0065	0.05	0.003139	0.01034
s2b1	0.3334	0.2302	4	1.45	0.2211	0.05	-0.3058	0.9726
s2b3	0.04374	0.03016	4	1.45	0.2206	0.05	-0.04000	0.1275