STAT436 **Homework 14** **Due Thursday 04/10 at 2.30pm**

Directions: 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. In this exercise, we reanalyze the Sitka89 data from Homework #13. This is done using SAS in the attached output, which uses one run of PROC NLIN and four runs of PROC NLMIXED. The NLIN and the first NLMIXED are run "by treatment," so each produces two sub-outputs; the respective outputs are identified by a corresponding title.
   (a) Identify the underlying assumptions and model that is fit in this first NLMIXED (Output B).
   (b) Focusing on the NLMIXED in Output B, make a 2x4 table of parameter estimates and comment on which parameters seem close (guesses can be made using approx. CIs) for the two treatments.
   (c) It turns out that Outputs B and C are virtually identical, noting that the -2LL's for Output B sum to [(-206.1) + (-492.8)], which is almost that of Output C (-697.2). In the second NLMIXED (Output C), what is the role of the "add" terms (i.e., as in "th1add", "th2add", etc.)?
   (d) Of the Outputs A, C, D, and E, which are special cases of others (that is, which are nested) – and which models are they nested in? List all nestings.
   (e) Clearly indicated what hypothesis can be tested by comparing Outputs C and D? Clearly list the hypotheses, test statistic, p-value and your conclusion. Which parameters are random in these models? Do these models assume that these random parameters vary by different or by the same amounts?
   (f) Answer the questions in part (e) but now comparing outputs D and E.
   (g) Using the model you feel best describes these data, summarize the data with your model, and comment on the model fit.
   (h) In comparing the NLIN output and the last NLMIXED, comment on the differences in the estimated variability associated with the LD_{50} parameters for each of these models.

2. A patient swallows a tablet of Zantac, which enters the patient's gut, and begins entering the patient's bloodstream at time t = 0. Blood samples are then taken every half-hour until hour 16, and the concentration of Zantac in the patient's serum are recorded; these data are analyzed in the SAS program in the Appendix.
   (a) Describe the model(s) that are being fit in the NLIN and the NLMIXED, and the associated assumptions. Are they fitting the same model? What are the model parameters, and the roles of these for each model.
   (b) The residual plot after the NLIN (plotting residuals versus lagged-residuals) highlights a problem with one of the implicit NLIN assumptions. Which assumption is it, and what is wrong? What are the usual ramifications of the violation of this (or these) assumption(s)?
   (c) Explain what is being done in the IML procedure, focusing on which function is being minimized in the 'NEG2LLA' and the 'NEG2LLB' functions. The latter function introduces an additional parameter; which one is it and what is its role? Is it "significant"? (Listing your null and alternative hypotheses, report the relevant test statistic, distribution, degrees of freedom and p-value.) Give reliable 90%, 95% and 99% CIs for this parameter.

First Exercise Output A – NLIN

```
proc sort data=sitka89; by treat; run;
proc nlin data=sitka89;
  parms th1=6.6 th2=5.5 th3=100 th4=3.5;
  t=(timea/th3)**th4; den=1+t;
  mean=th1+(th2-th1)/den;
  model size=mean;
  by treat;
run;
```

```
---------------------------------- treat=control ----------------------------------

                            The NLIN Procedure
                          Dependent Variable size

                                      Sum of        Mean                Approx
Source                        DF     Squares      Square    F Value     Pr > F
Model                          3     13.4823      4.4941      12.48     <.0001
Error                        196     70.5751      0.3601
Corrected Total              199     84.0574

                                 Approx
Parameter       Estimate      Std Error     Approximate 95% Confidence Limits
th1               6.5229         0.1270       6.2725       6.7733
th2               5.8491         0.0938       5.6641       6.0341
th3             76.1525         15.6448      45.2983        107.0
th4               3.6800         2.7532      -1.7497       9.1097


---------------------------------- treat=ozone ----------------------------------

                            The NLIN Procedure

                                      Sum of        Mean                Approx
Source                        DF     Squares      Square    F Value     Pr > F
Model                          3     28.0164      9.3388      21.75     <.0001
Error                        428       183.8      0.4294
Corrected Total              431       211.8

                                 Approx
Parameter       Estimate      Std Error     Approximate 95% Confidence Limits
th1               6.1481         0.0769       5.9969       6.2992
th2               5.5122         0.0665       5.3814       5.6429
th3             76.6292         10.4804      56.0294      97.2290
th4               4.3338         2.3368      -0.2594       8.9269
```

First Exercise Output B – NLMIXED #1

```
proc sort data=sitka89; by treat; run;
proc nlmixed data=sitka89;
  parms th1=6.6 th2=5.5 th3=100 th4=3.5 se2=.5 sa2=.1;
```

```
  th1a=th1+a; th2a=th2+a;
  t=(timea/th3)**th4; den=1+t;
  mean=th1a+(th2a-th1a)/den;
  model size~normal(mean,se2);
  random a~normal(0,sa2) subject=tree;
  by treat;
run;
```

The NLMIXED Procedure

Specifications

Data Set	WORK.SITKA89
Dependent Variable	size
Distribution for Dependent Variable	Normal
Random Effects	a
Distribution for Random Effects	Normal
Subject Variable	tree
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

---------- treat=control ----------

Dimensions

Observations Used	200
Observations Not Used	0
Total Observations	200
Subjects	25
Max Obs Per Subject	8
Parameters	6
Quadrature Points	1

Fit Statistics

-2 Log Likelihood	-206.1
AIC (smaller is better)	-194.1
AICC (smaller is better)	-193.7
BIC (smaller is better)	-186.8

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	6.5229	0.1188	24	54.92	<.0001	0.05	6.2778	6.7680
th2	5.8491	0.1181	24	49.51	<.0001	0.05	5.6053	6.0929
th3	76.1522	2.6475	24	28.76	<.0001	0.05	70.6881	81.6163
th4	3.6801	0.4346	24	8.47	<.0001	0.05	2.7831	4.5770
se2	0.01040	0.001112	24	9.35	<.0001	0.05	0.008108	0.01270
sa2	0.3425	0.09724	24	3.52	0.0017	0.05	0.1418	0.5432

---------- treat=ozone ----------

Dimensions

Observations Used	432
Observations Not Used	0
Total Observations	432
Subjects	54
Max Obs Per Subject	8
Parameters	6
Quadrature Points	1

```
                          Fit Statistics
                  -2 Log Likelihood                -492.8
                  AIC (smaller is better)          -480.8
                  AICC (smaller is better)         -480.6
                  BIC (smaller is better)          -468.9

                       Parameter Estimates
                    Standard
Parameter  Estimate     Error    DF  t Value  Pr > |t|   Alpha     Lower     Upper
th1          6.1481   0.08849    53    69.48    <.0001    0.05    5.9706    6.3256
th2          5.5122   0.08835    53    62.39    <.0001    0.05    5.3350    5.6894
th3         76.6290    1.5106    53    50.73    <.0001    0.05   73.5990   79.6589
th4          4.3339    0.3317    53    13.07    <.0001    0.05    3.6686    4.9992
se2        0.008920  0.000649    53    13.75    <.0001    0.05  0.007619   0.01022
sa2          0.4165   0.08037    53     5.18    <.0001    0.05    0.2553    0.5777
```

First Exercise Output C – NLMIXED #2

```
proc nlmixed data=sitka89;
  parms th1=6.6 th1add=-0.4 th2=5.8 th2add=-0.3 th3=76
        th3add=0.5 th4=3.7 th4add=0.6 se2=.01 sa2=.4;
  th1a=th1+th1add*dumoz+a;
  th2a=th2+th2add*dumoz+a;
  th3a=th3+th3add*dumoz;
  th4a=th4+th4add*dumoz;
  t=(timea/th3a)**th4a; den=1+t;
  mean=th1a+(th2a-th1a)/den;
  model size~normal(mean,se2);
  random a~normal(0,sa2) subject=tree;
run;
```

```
                        The NLMIXED Procedure

                           Dimensions
                  Observations Used                632
                  Observations Not Used              0
                  Total Observations               632
                  Subjects                          79
                  Max Obs Per Subject                8
                  Parameters                        10
                  Quadrature Points                  1

                          Fit Statistics
                  -2 Log Likelihood                -697.2
                  AIC (smaller is better)          -677.2
                  AICC (smaller is better)         -676.9
                  BIC (smaller is better)          -653.5

                       Parameter Estimates
                    Standard
Parameter  Estimate     Error    DF  t Value  Pr > |t|   Alpha     Lower     Upper
th1          6.5226    0.1268    78    51.43    <.0001    0.05    6.2701    6.7752
th1add      -0.3745    0.1533    78    -2.44    0.0168    0.05   -0.6796  -0.06937
th2          5.8489    0.1263    78    46.31    <.0001    0.05    5.5975    6.1004
```

th2add	-0.3367	0.1527	78	-2.20	0.0304	0.05	-0.6408	-0.03263
th3	76.0886	2.5121	78	30.29	<.0001	0.05	71.0875	81.0898
th3add	0.5558	2.9519	78	0.19	0.8511	0.05	-5.3210	6.4325
th4	3.6806	0.4125	78	8.92	<.0001	0.05	2.8594	4.5019
th4add	0.6532	0.5348	78	1.22	0.2256	0.05	-0.4115	1.7180
se2	0.009389	0.000565	78	16.63	<.0001	0.05	0.008265	0.01051
sa2	0.3931	0.06273	78	6.27	<.0001	0.05	0.2682	0.5180

First Exercise Output D – NLMIXED #3

```
proc nlmixed data=sitka89;
  parms th1=6.6 th1add=-0.4 th2=5.8 th2add=-0.3 th3=76
        th4=3.7 se2=.01 sa2=.4;
  th1a=th1+th1add*dumoz+a;
  th2a=th2+th2add*dumoz+a;
  t=(timea/th3)**th4; den=1+t;
  mean=th1a+(th2a-th1a)/den;
  model size~normal(mean,se2);
  random a~normal(0,sa2) subject=tree;
run;
```

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	-695.7
AIC (smaller is better)	-679.7
AICC (smaller is better)	-679.5
BIC (smaller is better)	-660.8

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	6.5124	0.1261	78	51.66	<.0001	0.05	6.2614	6.7633
th1add	-0.3594	0.1522	78	-2.36	0.0207	0.05	-0.6623	-0.05649
th2	5.8561	0.1261	78	46.45	<.0001	0.05	5.6051	6.1071
th2add	-0.3470	0.1524	78	-2.28	0.0255	0.05	-0.6503	-0.04365
th3	76.4949	1.3254	78	57.71	<.0001	0.05	73.8563	79.1336
th4	4.0971	0.2641	78	15.51	<.0001	0.05	3.5713	4.6230
se2	0.009415	0.000566	78	16.63	<.0001	0.05	0.008288	0.01054
sa2	0.3931	0.06273	78	6.27	<.0001	0.05	0.2682	0.5179

First Exercise Output E – NLMIXED #4

```
proc nlmixed data=sitka89;
  parms th1=6.6 th12add=-0.4 th2=5.8 th3=76 th4=3.7 se2=.01 sa2=.4;
  th1a=th1+th12add*dumoz+a;
  th2a=th2+th12add*dumoz+a;
  t=(timea/th3)**th4; den=1+t;
  mean=th1a+(th2a-th1a)/den;
  model size~normal(mean,se2);
  random a~normal(0,sa2) subject=tree;
run;
```

The NLMIXED Procedure

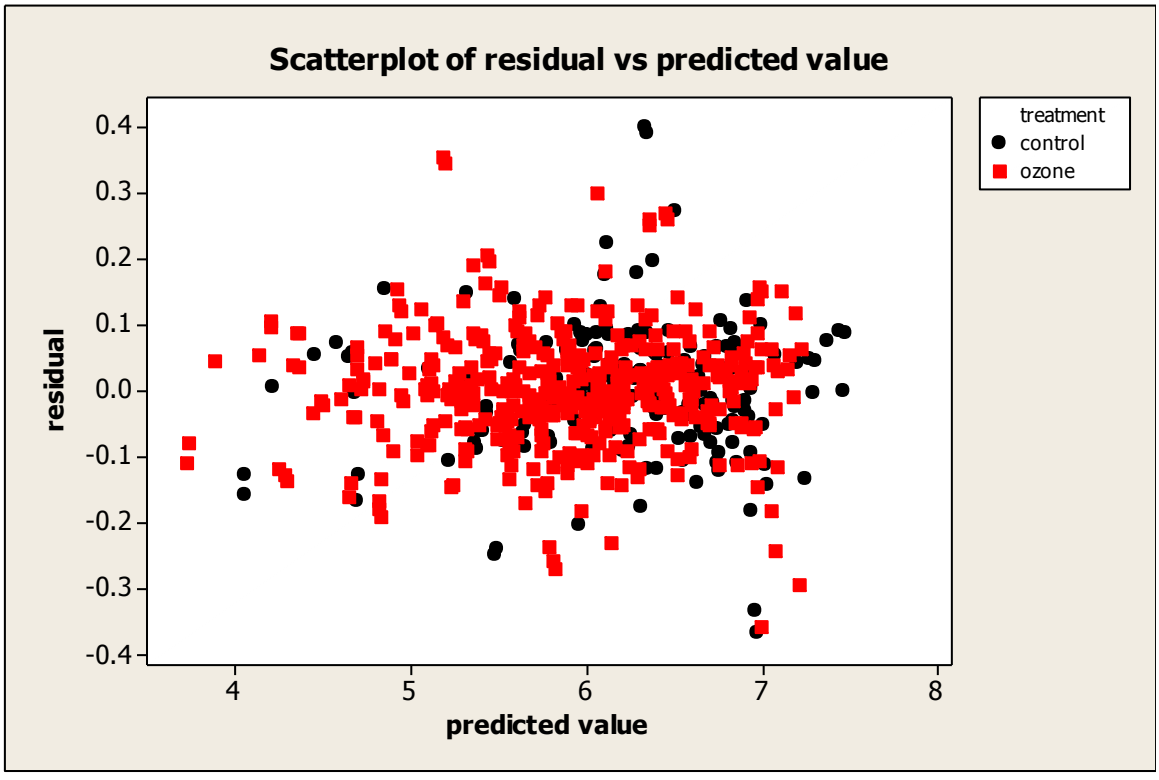
Fit Statistics

-2 Log Likelihood	-695.4
AIC (smaller is better)	-681.4
AICC (smaller is better)	-681.2
BIC (smaller is better)	-664.8

Parameter Estimates

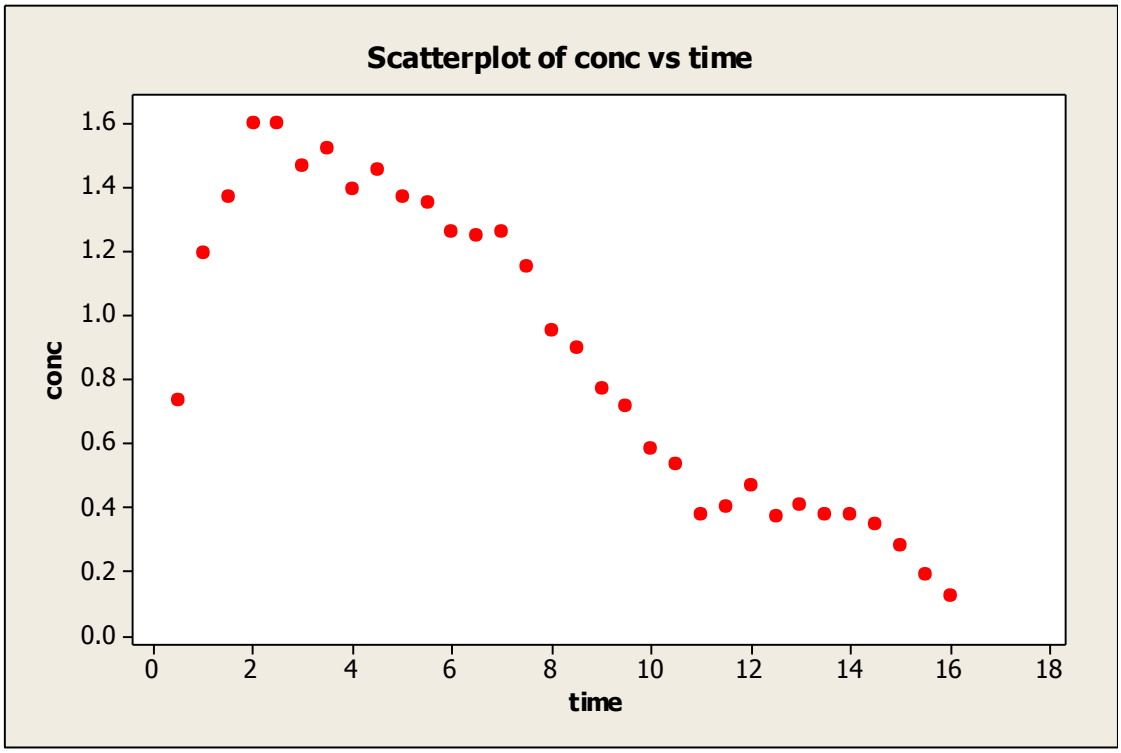
Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	6.5087	0.1259	78	51.70	<.0001	0.05	6.2581	6.7593
th12add	-0.3541	0.1519	78	-2.33	0.0223	0.05	-0.6565	-0.05173
th2	5.8610	0.1258	78	46.59	<.0001	0.05	5.6106	6.1114
th3	76.4968	1.3252	78	57.72	<.0001	0.05	73.8585	79.1350
th4	4.1000	0.2644	78	15.51	<.0001	0.05	3.5737	4.6263
se2	0.009421	0.000567	78	16.63	<.0001	0.05	0.008293	0.01055
sa2	0.3931	0.06273	78	6.27	<.0001	0.05	0.2682	0.5179

Residual Plot for NLMIXED #4



Second Exercise – Plot of Data

```
data one;
  do time=0.5 to 16 by 0.5;
    input conc @@; output;
  end; datalines;
0.73491 1.19450 1.36887 1.59890 1.59994 1.46580 1.51797 1.39002
1.45207 1.37051 1.35264 1.26148 1.24561 1.25900 1.15097 0.95105
0.89895 0.76953 0.71271 0.57851 0.53198 0.37821 0.40231 0.46940
0.36961 0.40307 0.37491 0.37525 0.34422 0.28192 0.18845 0.12393
;
proc plot;
  plot conc*time;
run;
```



Second Exercise – NLIN program and output

```
proc nlin;
  parms th1=0.4 th2=0.1 auc=16;
  ex1=exp(-th1*time); ex2=exp(-th2*time); diff=ex2-ex1;
  model conc=auc*(th1*th2/(th1-th2))*diff;
  output out=two r=r p=p;
run;
proc plot;
  plot r*p;
run;
```

The NLIN Procedure
Dependent Variable conc

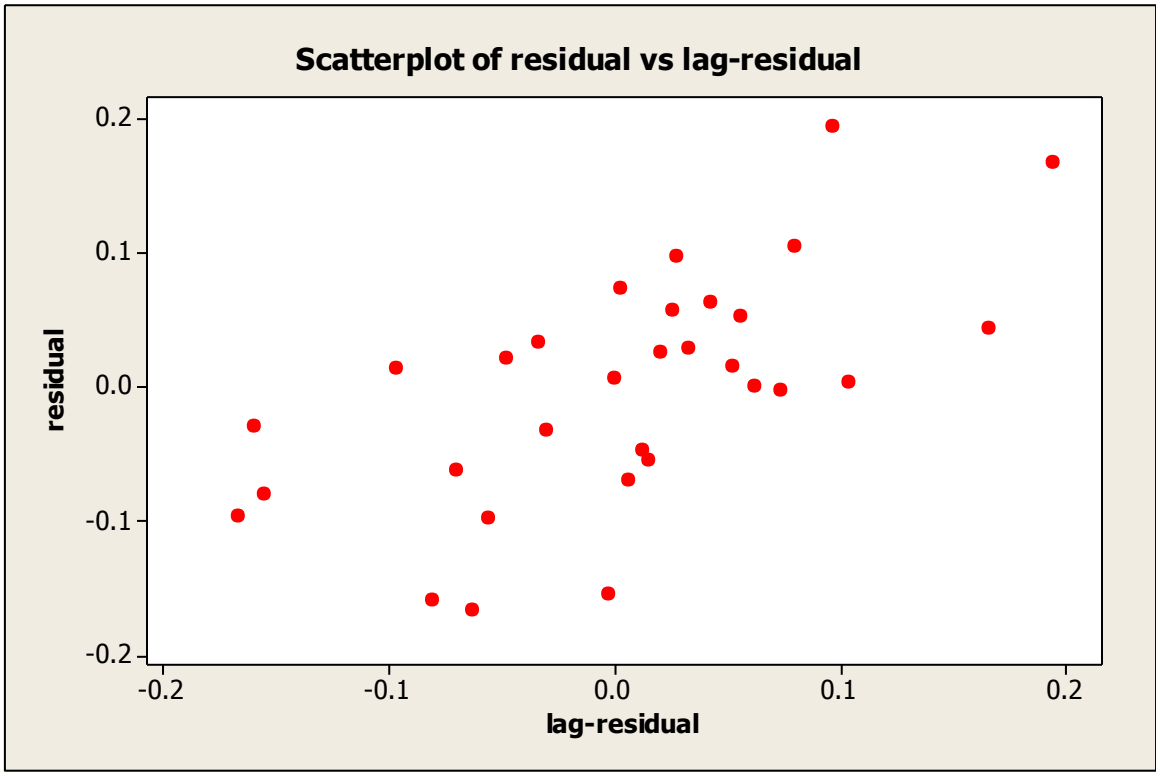
Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	3	31.7650	10.5883	1328.84	<.0001
Error	29	0.2311	0.00797		
Uncorrected Total	32	31.9960			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
th1	0.5709	0.0653	0.4373	0.7045
th2	0.1811	0.0153	0.1498	0.2123
auc	15.2610	0.4088	14.4249	16.0971

Approximate Correlation Matrix

	th1	th2	auc
th1	1.0000000	-0.9031835	0.6001821
th2	-0.9031835	1.0000000	-0.7733192
auc	0.6001821	-0.7733192	1.0000000

Plot of Residuals (from above NLIN) versus Lagged-Residuals



Second Exercise – NLMIXED program and output

```
proc nlmixed ecov data=one;
  parms th1=0.8 th2=0.05 phi=16 sig=0.06;
  ex1=exp(-th1*time); ex2=exp(-th2*time); diff=ex2-ex1;
  rec=1/(th1-th2); tmax=rec*log(th1/th2);
  cmax=th1*phi*exp(-th1*tmax);
  mean=th1*th2*phi*rec*diff; var=sig*sig;
  model conc~normal(mean,var);
  estimate 't-max' tmax;
  estimate 'c-max' cmax;
run;
```

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	-67.0
AIC (smaller is better)	-59.0
AICC (smaller is better)	-57.5
BIC (smaller is better)	-53.1

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	0.5709	0.06888	32	8.29	<.0001	0.05	0.4306	0.7112
th2	0.1811	0.01561	32	11.60	<.0001	0.05	0.1493	0.2129
phi	15.2610	0.3989	32	38.26	<.0001	0.05	14.4485	16.0735
sig	0.08498	0.01062	32	8.00	<.0001	0.05	0.06334	0.1066

Additional Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
t-max	2.9457	0.1233	32	23.89	<.0001	0.05	2.6945	3.1968
c-max	1.6211	0.02913	32	55.65	<.0001	0.05	1.5617	1.6804

Covariance Matrix of Additional Estimates

Row	Label	Cov1	Cov2
1	t-max	0.01520	-0.00076
2	c-max	-0.00076	0.000848

Second Exercise – IML program and output

```
proc iml;
  start neg2lla(th) global(xx,yy,nn);
    th1=th[1]; th2=th[2]; phi=th[3]; sig=th[4]; sig2=sig*sig;
    ex1=exp(-th1*xx); ex2=exp(-th2*xx); diff=ex2-ex1; rec=1/(th1-th2);
    eta=th1*th2*phi*rec*diff; res=yy-eta;
    thingy=nn*log(sig2)+(1/sig2)*t(res)*res;
    return(thingy);
  finish neg2lla;
```

```
  start neg2llb(th) global(xx,yy,nn); nn2=nn-1; on2=j(nn2,1);
    th1=th[1]; th2=th[2]; phi=th[3]; sig2=th[4]; rho=th[5];
    ex1=exp(-th1*xx); ex2=exp(-th2*xx); diff=ex2-ex1; rec=1/(th1-th2);
    eta=th1*th2*phi*rec*diff; res=yy-eta; rest=res[2:nn,]; ress=res[1:(nn-1),];
    xxt=xx[2:nn,]; xxs=xx[1:(nn-1),]; xxp=xxt-xxs; rhop=rho##(xxp);
    respp=rest-rhop#ress; rhop2=rho##(2*xxp); logrp=log(1-rhop2);
    wt=sqrt(1/(1-rhop2)); resp=wt#respp;
    thingy=nn*log(sig2)+t(logrp)*on2+(1/sig2)*( res[1]*res[1]+t(resp)*resp );
    return(thingy);
  finish neg2llb;

  xx={0.5,1,1.5,2,2.5,3,3.5,4,4.5,5,5.5,6,6.5,7,7.5,8,8.5,9,9.5,
       10,10.5,11,11.5,12,12.5,13,13.5,14,14.5,15,15.5,16};
  yy={0.73491,1.19450,1.36887,1.59890,1.59994,1.46580,1.51797,1.39002,
      1.45207,1.37051,1.35264,1.26148,1.24561,1.25900,1.15097,0.95105,
      0.89895,0.76953,0.71271,0.57851,0.53198,0.37821,0.40231,0.46940,
      0.36961,0.40307,0.37491,0.37525,0.34422,0.28192,0.18845,0.12393};
  nn=nrow(xx); opt={.,0};
  th0a={0.7 0.1 10 1};
  cona={0.001 0.000001 1 0.0001,
          .       .    .    .  };
  th0b={0.7 0.1 10 1 0.7};
  conb={0.001 0.000001 1 0.0001 0.0001,
          .       .    .    .   0.9999};
  call nlptr(rc,thetaa,"neg2lla",th0a,opt,cona);
     minn2lla=neg2lla(thetaa); print thetaa minn2lla;
  call nlptr(rc,thetab,"neg2llb",th0b,opt,conb);
     minn2llb=neg2llb(thetab); print thetab minn2llb;
  neg2llmin=neg2llb(thetab); ans=j(50,5);
  do jj=1 to 50;
    ans[jj,3]=neg2llmin+2.71;
    ans[jj,4]=neg2llmin+3.84;
    ans[jj,5]=neg2llmin+6.63;
    star=0.05; fini=0.97; rho=star+(fini-star)*(jj-1)/49; ans[jj,1]=rho;
    th0b=thetab; opt={.,0};
    th0b[5]=rho; conb[1,5]=rho; conb[2,5]=rho;
    call nlptr(rc,thtil,"neg2llb",th0b,opt,conb);
    ans[jj,2]=neg2llb(thtil);
  end;
  create dset from ans[colname={rho negll cut90 cut95 cut99}];
  append from ans;
quit;
data final; set dset;
  y=negll;  zz=1; output;
  y=cut90;  zz=2; output;
  y=cut95;  zz=3; output;
  y=cut99;  zz=4; output;
axis1 w=2 minor=none label=(a=90 font=swiss 'Profile Likelihood');
axis2 w=2 minor=none label=(font=swiss 'RHO - AR1 Parameter');
symbol1 i=join v=point w=5 l=1 c=blue;
symbol2 i=join v=point w=3 l=2 c=blue;
symbol3 i=join v=point w=3 l=3 c=blue;
symbol4 i=join v=point w=3 l=4 c=blue;
proc gplot data=final;
  plot y*rho=zz / fr vaxis=axis1 nolegend haxis=axis2;
run;
```

```
THETAA                                                    MINN2LLA
0.1810802 0.5708944 15.260993 0.0849769                   -125.784

THETAB                                                    MINN2LLB
0.6247655 0.1728786 15.344363 0.0073746 0.4185682         -141.9209
```

Second Exercise – Profile Likelihood Plot for Parameter ϕ