STAT436 **Homework 10** **Due Tuesday 03/11 at 2.30pm**

<u>Directions</u>: 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 "Calibration and assay development using the four-parameter logistic curve" (*Chem. Intell. Lab. Systems*, 1993, p.97), O'Connell *et al* fit the four-parameter log-logistic (LL4) model function to their radioimmunoassay (RIA) data. The data are analyzed in SAS in the Appendix using **proc nlin** and then **three** runs of **proc nlmixed**.

(a) Clearly list all the needed assumptions for the run of **proc nlin** and the three runs of **proc nlmixed** <u>in the context of this exercise/situation</u>.

(b) Looking at the residual plot of the **proc nlin** fit in Output A (p.3), comment on whether all necessary assumptions appear to be met. (Incidentally, note that the **proc nlmixed** fit in Output C fits this same ***homoskedastic*** – i.e., constant variance – nonlinear model as in this **proc nlin**).

(c) In Output B, researchers are trying to obtain a good model for the model variance for these data. Explain what is being done in the **proc nlmixed** in Output D: what is the 'model' and what are the roles of the new parameter(s). Perform a **<u>likelihood-based test</u>** of whether the extra parameter (rho = ρ) in the variance is need, writing out your hypotheses, test statistic, degrees of freedom, p-value, and clear conclusion. Does this conclusion seem sensible in light of the residual plot on p.3?

(d) It turns out that the **proc nlmixed** in Output E involves another – more appropriate but more complicated – way of modeling the variance for these data. Comparing Outputs C and E, perform a **<u>likelihood-based hypothesis test</u>** testing for homoskedasticity, again writing out your hypotheses, test statistic, degrees of freedom, p-value, and conclusion.

(e) Finally, compare the parameter estimates for the model parameters (the θs) and – more importantly – the associated standard errors for Outputs C and E – what has changed?

2. Data are graphed and examined in Output 2 using two runs of **proc nlin**.

(a) Clearly list all the needed assumptions for the first **proc nlin** analysis.

(b) Write down the model function being fit in Output A. What is the precise relevance/interpretation of the parameter named 'phi' (denoted ϕ)?

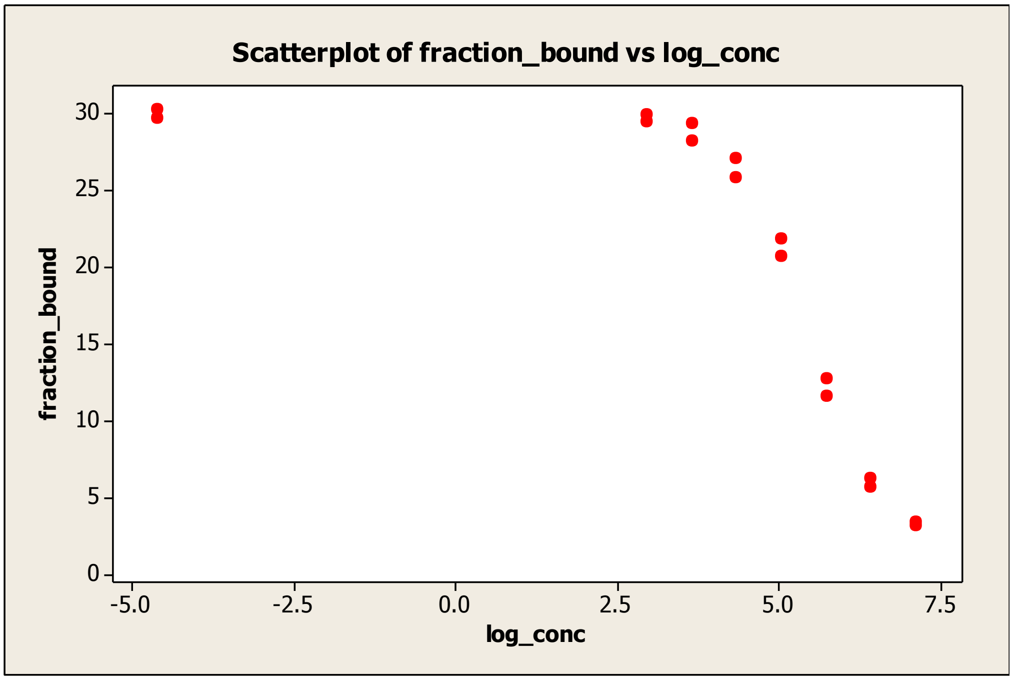
(c) Write down the estimate of 'phi'. Also, perform a two-sided (WALD) t-test that the true value of 'phi' is equal to **-0.40** (using α = 5%), and write down the 95 % Wald Confidence Interval (WCI) for this parameter. For the test, give the hypotheses, test statistic, p-value and degrees of freedom.

(d) Repeat part (c) using the likelihood test. Remember to give the hypotheses, test statistic, p-value and degrees of freedom.

(e) It turns out that the true 95% confidence interval for 'phi' for these data is very different from the reported WCI. In your opinion, is this true CI shifted to the right or to the left of the WCI? Clearly and succinctly explain your answer.

Homework 10 Appendix

Exercise 1 Graph and Data Listing



conc	fraction_bound	log_conc
0.0	30.16	-4.60517
0.0	29.58	-4.60517
19.4	29.87	2.96579
19.4	29.43	2.96579
38.8	28.19	3.65868
38.8	29.33	3.65868
77.5	26.96	4.35041
77.5	25.72	4.35041
155.0	21.82	5.04349
155.0	20.59	5.04349
310.0	12.62	5.73660
310.0	11.57	5.73660
620.0	5.56	6.42974
620.0	6.17	6.42974
1240.0	3.33	7.12287
1240.0	3.07	7.12287

Exercise 1 SAS Program A

```
proc nlin data=one;
  parms th1=30 th2=0 th3=300 th4=2;
  if conc=0 then do;
    mean=th1;
```

```
  end;
  else do;
    t=(conc/th3)**th4;
    mean=th2+(th1-th2)/(1+t);
  end;
  model fraction_bound=mean;
  output out=two r=residual p=predicted_value;
run;
proc print noobs;
  var residual predicted_value;
run;
```

Exercise 1 SAS Output A

The NLIN Procedure
Dependent Variable fraction_bound

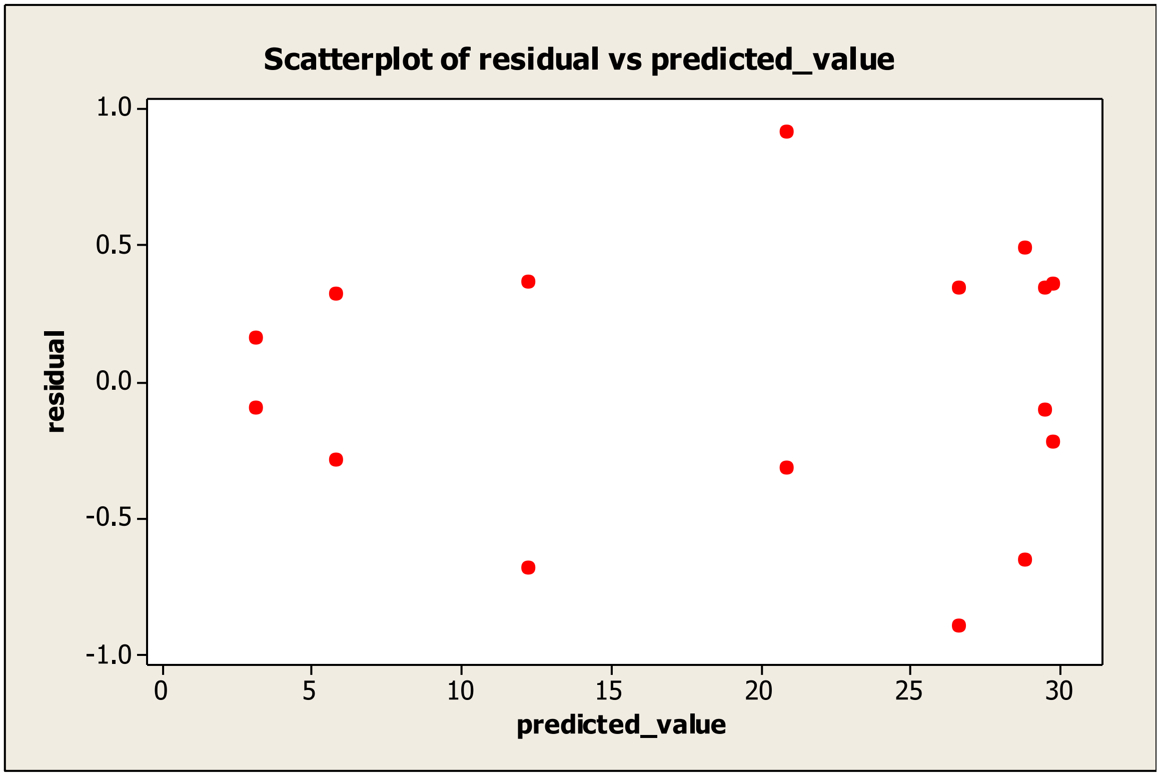
Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Model	3	1704.2	568.1	1869.08	<.0001
Error	12	3.6471	0.3039		
Corrected Total	15	1707.8			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
th1	29.8033	0.2658	29.2242	30.3823
th2	1.9929	0.5667	0.7582	3.2276
th3	232.4	8.6272	213.6	251.2
th4	1.8619	0.1122	1.6174	2.1065

Listing of Residuals and Predicted Values

residual	predicted_value
0.35675	29.8033
-0.22325	29.8033
0.33705	29.5329
-0.10295	29.5329
-0.65525	28.8452
0.48475	28.8452
0.34247	26.6175
-0.89753	26.6175
0.91179	20.9082
-0.31821	20.9082
0.36251	12.2575
-0.68749	12.2575
-0.28828	5.8483
0.32172	5.8483
0.15796	3.1720
-0.10204	3.1720

Residual Plot

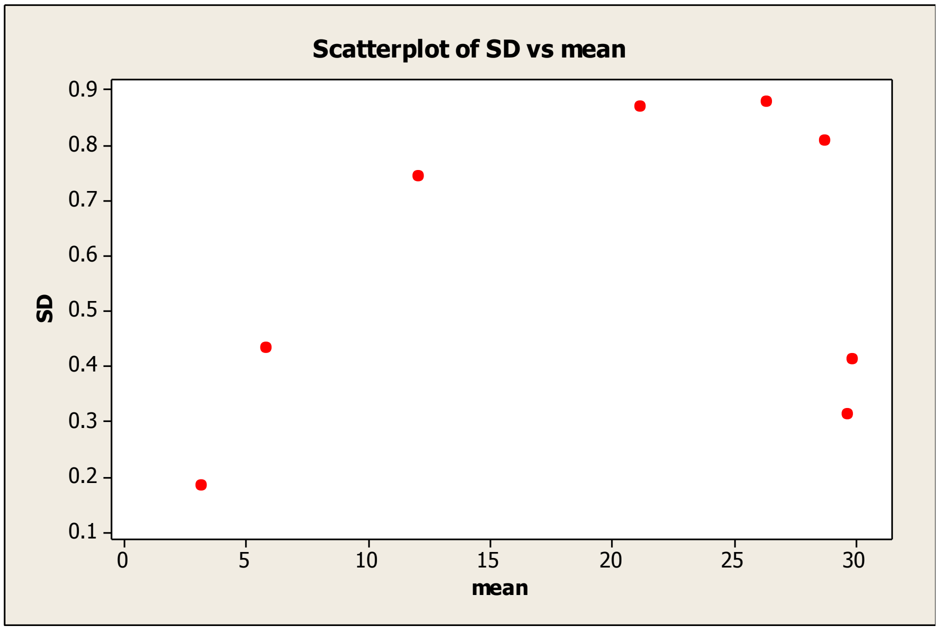


Exercise 1 SAS Program B

```
proc sort data=one;
  by conc;
run;
proc means noprint;
  var fraction_bound;
  by conc;
  output out=three mean=mean stddev=sd;
run;
data three; set three;
  drop _TYPE_ _FREQ_;
proc print; run;
```

Exercise 1 SAS Output B

Obs	conc	mean	sd
1	0.0	29.870	0.41012
2	19.4	29.650	0.31113
3	38.8	28.760	0.80610
4	77.5	26.340	0.87681
5	155.0	21.205	0.86974
6	310.0	12.095	0.74246
7	620.0	5.865	0.43134
8	1240.0	3.200	0.18385



Exercise 1 SAS Program and Output C

```
proc nlmixed data=one;
  parms th1=30 th2=0 th3=300 th4=2 sigma=1;
  if conc=0 then do;
    mean=th1;
  end;
  else do;
    t=(conc/th3)**th4; den=1+t;
    mean=th2+(th1-th2)/den;
  end;
  var=sigma*sigma;
  model fraction_bound ~ normal(mean,var);
run;
```

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	21.7
AIC (smaller is better)	31.7
AICC (smaller is better)	37.7
BIC (smaller is better)	35.6

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	29.8033	0.2329	16	127.97	<.0001	0.05	29.3095	30.2970
th2	1.9929	0.4927	16	4.04	0.0009	0.05	0.9484	3.0374
th3	232.44	7.4175	16	31.34	<.0001	0.05	216.72	248.17
th4	1.8619	0.09942	16	18.73	<.0001	0.05	1.6512	2.0727
sigma	0.4774	0.08440	16	5.66	<.0001	0.05	0.2985	0.6564

Exercise 1 SAS Program and Output D

```
proc nlmixed data=one;
  parms th1=30 th2=0 th3=300 th4=2 sigma=.01 rho=0;
  if conc=0 then do;
    mean=th1;
  end;
  else do;
    t=(conc/th3)**th4; den=1+t;
    mean=th2+(th1-th2)/den;
  end;
  var=sigma*sigma*(mean)**rho;
  model fraction_bound ~ normal(mean,var);
run;
```

The NLMIXED Procedure

Fit Statistics

-2 Log Likelihood	19.4
AIC (smaller is better)	31.4
AICC (smaller is better)	40.7
BIC (smaller is better)	36.0

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	29.7840	0.2777	16	107.24	<.0001	0.05	29.1952	30.3728
th2	2.0618	0.2811	16	7.33	<.0001	0.05	1.4659	2.6577
th3	231.74	5.6829	16	40.78	<.0001	0.05	219.69	243.79
th4	1.8784	0.08705	16	21.58	<.0001	0.05	1.6939	2.0629
sigma	0.1223	0.09460	16	1.29	0.2144	0.05	-0.07822	0.3229
rho	0.9410	0.5501	16	1.71	0.1065	0.05	-0.2252	2.1071

Exercise 1 SAS Program and Output E

```
proc nlmixed data=one;
  parms th1=30 th2=2 th3=250 th4=2 sigma=.01 rho1=5 rho2=4;
  if conc=0 then do;
    mean=th1;
  end;
  else do;
    t=(conc/th3)**th4; den=1+t;
    mean=th2+(th1-th2)/den;
  end;
  var=.00001*sigma*sigma*((mean)**rho1)*((1.1*th1-mean)**rho2);
  model fraction_bound ~ normal(mean,var);
run;
```

The NLMIXED Procedure

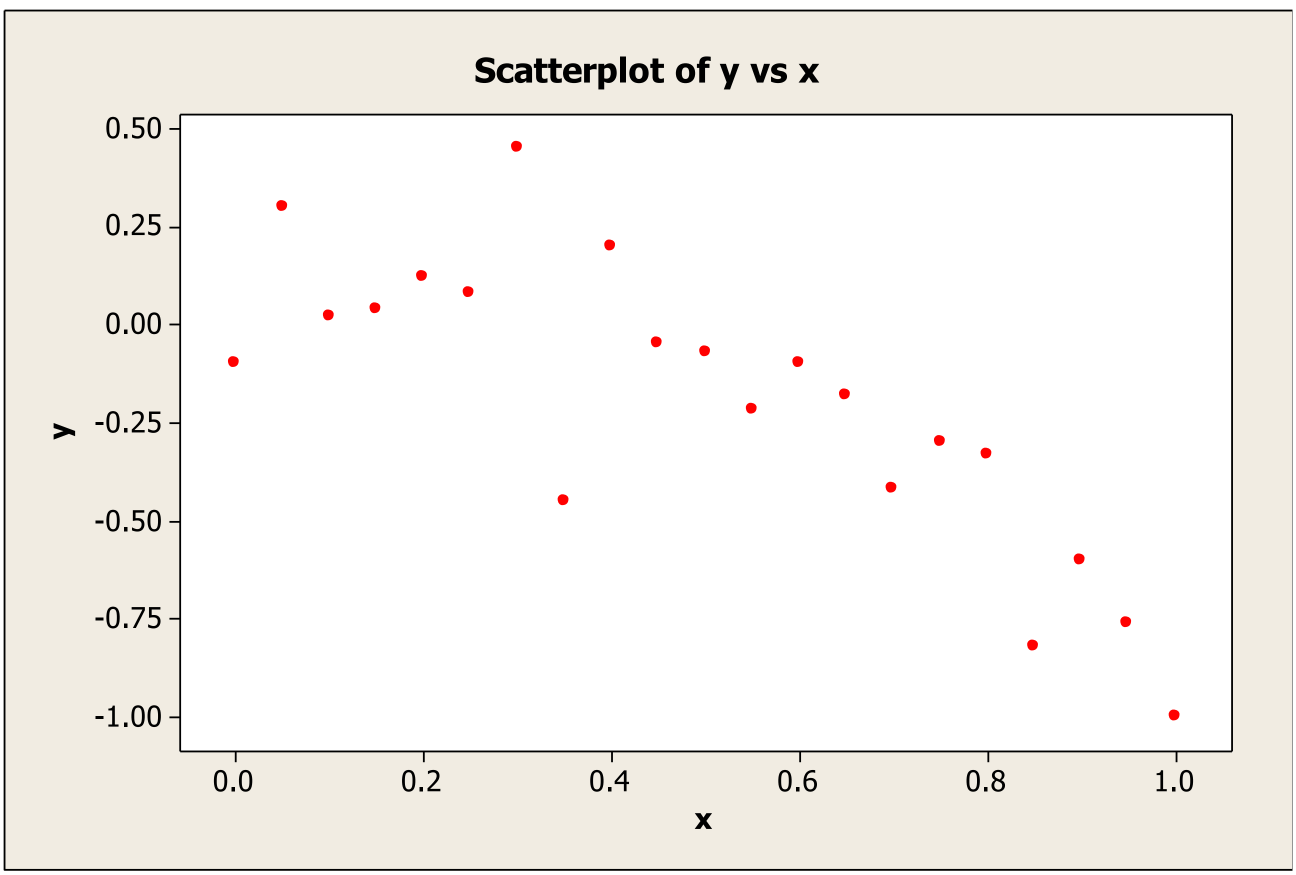
Fit Statistics

-2 Log Likelihood	15.4
AIC (smaller is better)	29.4
AICC (smaller is better)	43.4
BIC (smaller is better)	34.8

Parameter Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Lower	Upper
th1	29.8710	0.1771	16	168.64	<.0001	0.05	29.4955	30.2464
th2	1.9814	0.2295	16	8.63	<.0001	0.05	1.4948	2.4680
th3	230.04	6.7797	16	33.93	<.0001	0.05	215.67	244.42
th4	1.8332	0.08383	16	21.87	<.0001	0.05	1.6555	2.0109
sigma	0.2476	0.6171	16	0.40	0.6936	0.05	-1.0606	1.5557
rho1	2.9106	1.0545	16	2.76	0.0139	0.05	0.6751	5.1461
rho2	2.0161	0.9914	16	2.03	0.0589	0.05	-0.08554	4.1176

Exercise 2 Graph



Exercise 2 SAS Program and Output A

```
proc nlin;
  parms b0=0 b2=-.5 phi=.25;
  model y=b0-2*phi*b2*x+b2*x*x;
run;
```

```
                     The NLIN Procedure
                        Dependent Variable y
                        Method: Gauss-Newton

                          Iterative Phase
                                                         Sum of
         Iter          b0          b2         phi       Squares
            0           0     -0.5000      0.2500        2.3100
            1      0.0424     -1.5529      0.0508        1.9069
            2      0.0424     -1.5529      0.1859        0.6451
   NOTE: Convergence criterion met.

                         Estimation Summary
                 Method                   Gauss-Newton
                 Iterations                          2
                 Observations Read                  21
                 Observations Used                  21
                 Observations Missing                0

                                 Sum of        Mean               Approx
Source                    DF    Squares      Square    F Value    Pr > F
Model                      2     2.1702      1.0851      30.28    <.0001
Error                     18     0.6451      0.0358
Corrected Total           20     2.8153

                                 Approx
Parameter      Estimate      Std Error    Approximate 95% Confidence Limits
b0               0.0424         0.1130      -0.1950      0.2798
b2              -1.5529         0.5056      -2.6150     -0.4907
phi              0.1859         0.1113      -0.0480      0.4197

                      Approximate Correlation Matrix
                            b0                b2             phi
         b0          1.0000000         0.7084130      -0.8891903
         b2          0.7084130         1.0000000      -0.9188135
         phi        -0.8891903        -0.9188135       1.0000000
```

Exercise 2 SAS Program and Output B

```
proc nlin;
  parms b0=0 b2=-.5; phi=-0.40;
  model y=b0-2*phi*b2*x+b2*x*x;
run;
```

```
                    The NLIN Procedure
                   Dependent Variable y
                  Method: Gauss-Newton

                     Iterative Phase
                                              Sum of
          Iter          b0          b2       Squares
             0           0     -0.5000        1.4278
             1      0.2189     -0.5642        0.7852
   NOTE: Convergence criterion met.

                   Estimation Summary
             Method                  Gauss-Newton
             Iterations                         1
             Observations Read                 21
             Observations Used                 21
             Observations Missing               0

                             Sum of        Mean               Approx
Source                DF    Squares      Square    F Value    Pr > F
Model                  1     2.0301      2.0301      49.13    <.0001
Error                 19     0.7852      0.0413
Corrected Total       20     2.8153

                             Approx
Parameter      Estimate   Std Error    Approximate 95% Confidence Limits
b0               0.2189      0.0744       0.0633      0.3746
b2              -0.5642      0.0805      -0.7327     -0.3957

              Approximate Correlation Matrix
                          b0              b2
              b0       1.0000000      -0.8026837
              b2      -0.8026837       1.0000000
```