STAT436 **Homework 12** **Due Tuesday 4/1 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.

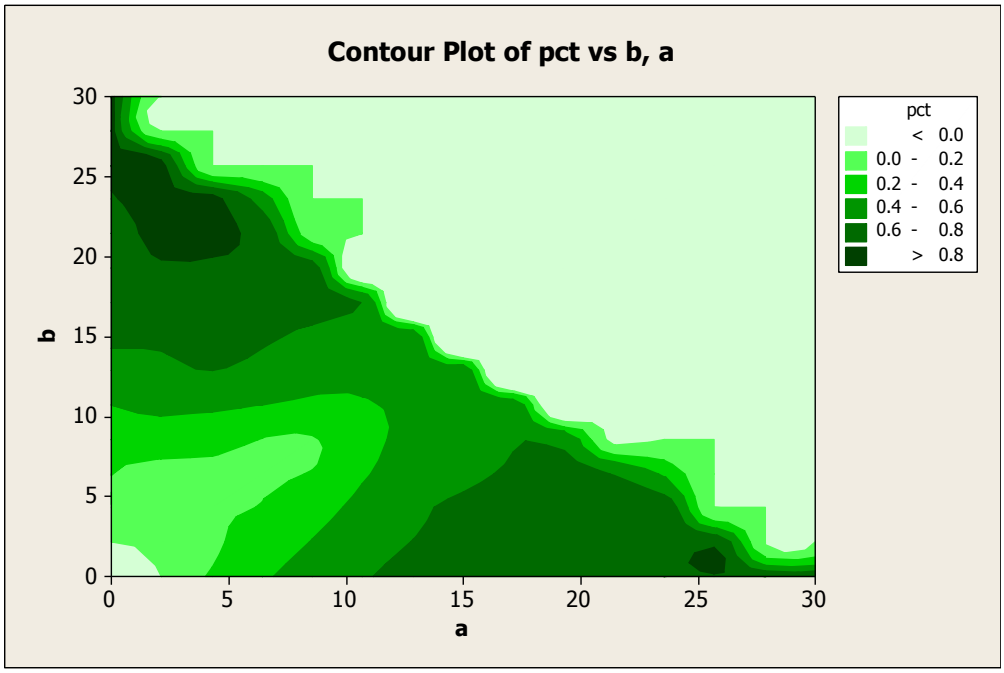
1. Data provided in Giltinan, Capizzi and Malani (1988), "Diagnostic tests for similar action of two compounds," *Applied Statistics*, resulted from an experiment designed to investigate how two insecticides (A and B) may act in combination. Of interest here is whether the insecticides interact to produce enhanced performance, a reduction in performance or neither. The data, corresponding to the number of dead insects corresponding to a fixed number of insects tested (usually 30) for 20 combinations of the insecticides are analyzed in the SAS program/output.
   (a) Comment on the **experimental design** used for this experiment (i.e., identify whether it is a ray, factorial or other design; how many interior rays if any are used and support points per ray, etc.). How many total experimental units (insects) were used for the study? Does the chosen design seem like an efficient way to use the experimental units, or would you have chosen another strategy? If so, what and why? If not, why (specifically) do you like this one?
   (b) Does this analysis suggest enhanced effect, a reduction in performance or neither of these? ***Using only the results of the first two PROC NLMIXED's for this part***, identifying the model that is being used here. Support your claim(s) about any interaction using two distinct test statistics and p-values; suggest which of these tests is preferred and why. Clearly give all assumptions.
   (c) In examining how well or otherwise the predicted ("Pred") and actual ("ratio") proportions coincide, comment on the adequacy of the above-used model for these data (see p.4).

2. Returning to the previous exercise and SAS program/output:
   (a) Describe (name) the model(s) that are being fit in the third and fourth NLMIXED's. Write down the hypotheses (null and alternative) that can be tested using these NLMIXED's, as well as the test statistic, p-value and your detailed conclusion. Clearly give all assumptions.
   (b) Connect the first and fourth NLMIXED's (is either a special case of the other?) and decide which is better; give a test statistic and p-value if possible and your clear conclusion.
   (c) Calculate the predicted values of the $LD_{50}$'s along the interior rays that the model given by the first NLMIXED and compare these with the $LD_{50}$'s for the fourth NLMIXED. Are they close?
   (d) In the general setting, give necessary and sufficient conditions for the Finney and Separate-ray models to be equivalent when the chosen design is a 2+3 ray design and when the logistic curve is fit as for this exercise.

## Homework 12 Appendix

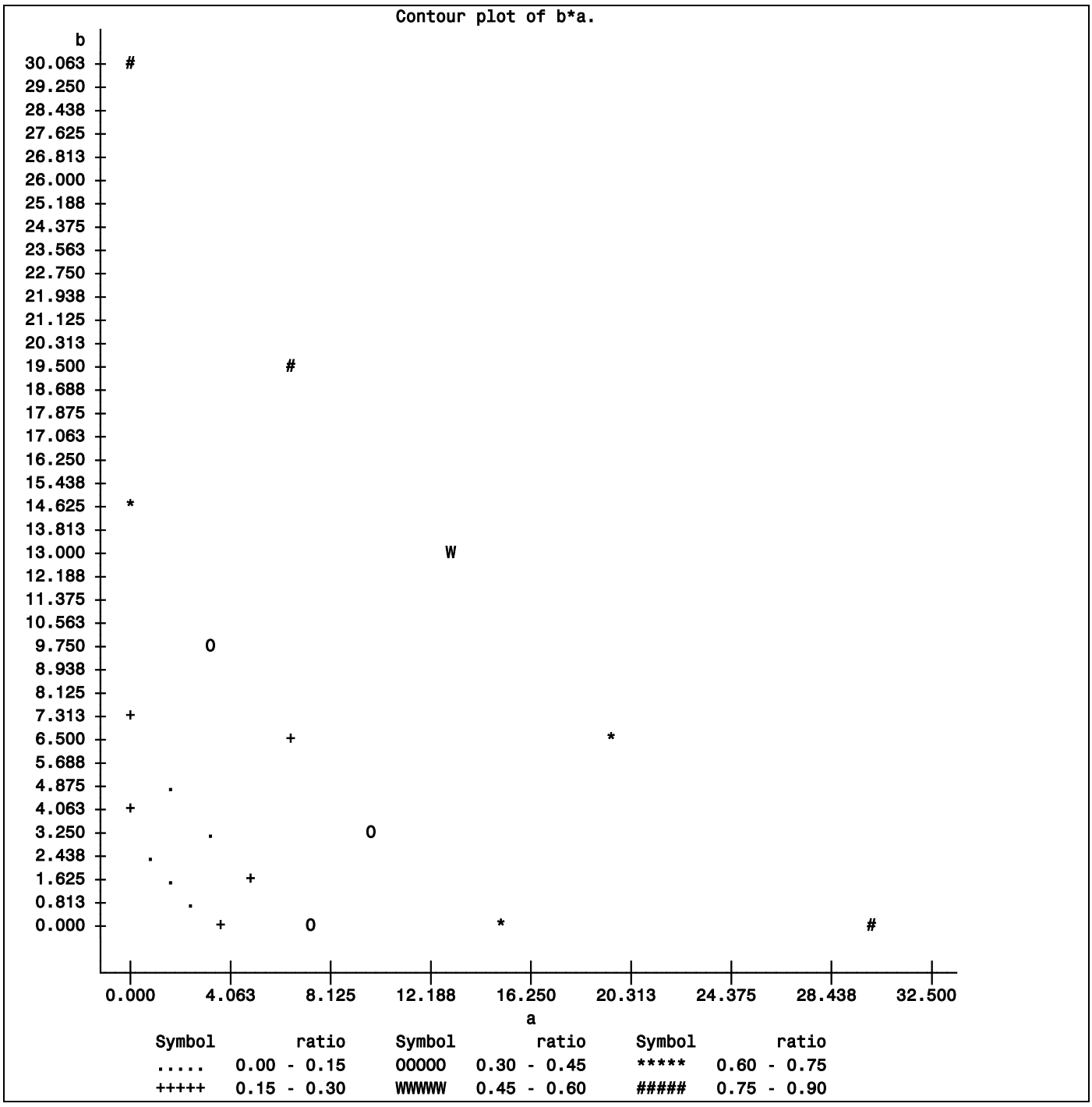
```
data one;
  input a b y n @@; ratio=y/n; datalines;
0.0000  30.0000     26 30 0.0000 15.0000 19 30 0.0000 7.5000 7 30 0.0000 3.7500 5 30
6.5000  19.5000 23 30 3.2500 9.7500 11 30 1.6250 4.8750 3 30 0.8125 2.4375 0 30
13.0000 13.0000 15 30 6.5000 6.5000 5 30 3.2500 3.2500 4 29 1.6250 1.6250 0 29
19.5000  6.5000 20 30 9.7500 3.2500 13 30 4.8750 1.6250 6 29 2.4375 0.8125 0 30
30.0000  0.0000 23 30 15.0000 0.0000 21 30 7.5000 0.0000 13 30 3.7500 0.0000 5 30
;
data one; set one;
  if a=0 then a=0.00000000000000000000000000000000000000000000001;
  if b=0 then b=0.00000000000000000000000000000000000000000000001; dd=a+b;
  ray=2*(_n_ le 4)+3*(_n_ le 8 and _n_ ge 5)+4*(_n_ le 12 and _n_ ge 9)
           +5*(_n_ le 16 and _n_ ge 13)+1*(_n_ ge 17);
proc print noobs data=one; run;
```

| a | b | y | n | ratio | dd | ray |
|---|---|---|---|---|---|---|
| 0.0000 | 30.0000 | 26 | 30 | 0.86667 | 30.00 | 2 |
| 0.0000 | 15.0000 | 19 | 30 | 0.63333 | 15.00 | 2 |
| 0.0000 | 7.5000 | 7 | 30 | 0.23333 | 7.50 | 2 |
| 0.0000 | 3.7500 | 5 | 30 | 0.16667 | 3.75 | 2 |
| 6.5000 | 19.5000 | 23 | 30 | 0.76667 | 26.00 | 3 |
| 3.2500 | 9.7500 | 11 | 30 | 0.36667 | 13.00 | 3 |
| 1.6250 | 4.8750 | 3 | 30 | 0.10000 | 6.50 | 3 |
| 0.8125 | 2.4375 | 0 | 30 | 0.00000 | 3.25 | 3 |
| 13.0000 | 13.0000 | 15 | 30 | 0.50000 | 26.00 | 4 |
| 6.5000 | 6.5000 | 5 | 30 | 0.16667 | 13.00 | 4 |
| 3.2500 | 3.2500 | 4 | 29 | 0.13793 | 6.50 | 4 |
| 1.6250 | 1.6250 | 0 | 29 | 0.00000 | 3.25 | 4 |
| 19.5000 | 6.5000 | 20 | 30 | 0.66667 | 26.00 | 5 |
| 9.7500 | 3.2500 | 13 | 30 | 0.43333 | 13.00 | 5 |
| 4.8750 | 1.6250 | 6 | 29 | 0.20690 | 6.50 | 5 |
| 2.4375 | 0.8125 | 0 | 30 | 0.00000 | 3.25 | 5 |
| 30.0000 | 0.0000 | 23 | 30 | 0.76667 | 30.00 | 1 |
| 15.0000 | 0.0000 | 21 | 30 | 0.70000 | 15.00 | 1 |
| 7.5000 | 0.0000 | 13 | 30 | 0.43333 | 7.50 | 1 |
| 3.7500 | 0.0000 | 5 | 30 | 0.16667 | 3.75 | 1 |




**Above contour plot is done using Minitab, and the one below using SAS**

```
proc plot;
  plot b*a=ratio/contour=6;
run;
```




**First NLMIXED**

```
proc nlmixed data=one;
  parms th2=10 th3=2 th4=1 th5=0;
  z=a+th4*b+th5*sqrt(th4*a*b);
  t=(z/th2)**th3; den=1+t; p=t/den;
  model y~binomial(n,p);
  predict p out=new;
run;
```

The NLMIXED Procedure

Specifications

| | |
|---|---|
| Data Set | WORK.ONE |
| Dependent Variable | y |
| Distribution for Dependent Variable | Binomial |
| Optimization Technique | Dual Quasi-Newton |
| Integration Method | None |

Dimensions

| | |
|---|---|
| Observations Used | 20 |
| Observations Not Used | 0 |
| Total Observations | 20 |
| Parameters | 4 |

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 80.6 |
| AIC (smaller is better) | 88.6 |
| AICC (smaller is better) | 91.3 |
| BIC (smaller is better) | 92.6 |

Parameter Estimates

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| th2 | 9.9531 | 1.0590 | 20 | 9.40 | <.0001 | 0.05 | 7.7441 | 12.1621 |
| th3 | 1.8018 | 0.1612 | 20 | 11.18 | <.0001 | 0.05 | 1.4655 | 2.1381 |
| th4 | 0.9002 | 0.1170 | 20 | 7.69 | <.0001 | 0.05 | 0.6562 | 1.1442 |
| th5 | -1.0349 | 0.1390 | 20 | -7.44 | <.0001 | 0.05 | -1.3248 | -0.7449 |

```
proc print data=new;
  title 'predicted and actual probabilities from First NLMixed'; run;
```

| Obs | a | b | y | n | ratio | dd | ray | Pred |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.0000 | 30.0000 | 26 | 30 | 0.86667 | 30.00 | 2 | 0.85796 |
| 2 | 0.0000 | 15.0000 | 19 | 30 | 0.63333 | 15.00 | 2 | 0.63403 |
| 3 | 0.0000 | 7.5000 | 7 | 30 | 0.23333 | 7.50 | 2 | 0.33196 |
| 4 | 0.0000 | 3.7500 | 5 | 30 | 0.16667 | 3.75 | 2 | 0.12474 |
| 5 | 6.5000 | 19.5000 | 23 | 30 | 0.76667 | 26.00 | 3 | 0.61802 |
| 6 | 3.2500 | 9.7500 | 11 | 30 | 0.36667 | 13.00 | 3 | 0.31696 |
| 7 | 1.6250 | 4.8750 | 3 | 30 | 0.10000 | 6.50 | 3 | 0.11746 |
| 8 | 0.8125 | 2.4375 | 0 | 30 | 0.00000 | 3.25 | 3 | 0.03677 |
| 9 | 13.0000 | 13.0000 | 15 | 30 | 0.50000 | 26.00 | 4 | 0.58120 |
| 10 | 6.5000 | 6.5000 | 5 | 30 | 0.16667 | 13.00 | 4 | 0.28471 |
| 11 | 3.2500 | 3.2500 | 4 | 29 | 0.13793 | 6.50 | 4 | 0.10247 |
| 12 | 1.6250 | 1.6250 | 0 | 29 | 0.00000 | 3.25 | 4 | 0.03171 |
| 13 | 19.5000 | 6.5000 | 20 | 30 | 0.66667 | 26.00 | 5 | 0.65759 |
| 14 | 9.7500 | 3.2500 | 13 | 30 | 0.43333 | 13.00 | 5 | 0.35518 |
| 15 | 4.8750 | 1.6250 | 6 | 29 | 0.20690 | 6.50 | 5 | 0.13643 |
| 16 | 2.4375 | 0.8125 | 0 | 30 | 0.00000 | 3.25 | 5 | 0.04335 |
| 17 | 30.0000 | 0.0000 | 23 | 30 | 0.76667 | 30.00 | 1 | 0.87952 |
| 18 | 15.0000 | 0.0000 | 21 | 30 | 0.70000 | 15.00 | 1 | 0.67678 |
| 19 | 7.5000 | 0.0000 | 13 | 30 | 0.43333 | 7.50 | 1 | 0.37522 |
| 20 | 3.7500 | 0.0000 | 5 | 30 | 0.16667 | 3.75 | 1 | 0.14694 |

## Second NLMIXED

```
proc nlmixed data=one;
  parms th2=10 th3=2 th4=1; th5=0;
  z=a+th4*b+th5*sqrt(th4*a*b);
  t=(z/th2)**th3; den=1+t; p=t/den;
  model y~binomial(n,p);
run;
```

The NLMIXED Procedure

Specifications

| | |
|---|---|
| Data Set | WORK.ONE |
| Dependent Variable | y |
| Distribution for Dependent Variable | Binomial |
| Optimization Technique | Dual Quasi-Newton |
| Integration Method | None |

Dimensions

| | |
|---|---|
| Observations Used | 20 |
| Observations Not Used | 0 |
| Total Observations | 20 |
| Parameters | 3 |

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 110.9 |
| AIC (smaller is better) | 116.9 |
| AICC (smaller is better) | 118.4 |
| BIC (smaller is better) | 119.9 |

Parameter Estimates

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| th2 | 14.4553 | 1.3289 | 20 | 10.88 | <.0001 | 0.05 | 11.6833 | 17.2272 |
| th3 | 1.7352 | 0.1522 | 20 | 11.40 | <.0001 | 0.05 | 1.4177 | 2.0527 |
| th4 | 0.8962 | 0.1291 | 20 | 6.94 | <.0001 | 0.05 | 0.6268 | 1.1656 |

## Third NLMIXED

```
proc nlmixed data=one;
  parms th21=10 th22=15 k3=1 k4=1 k5=1 th31=2 th32=2 th33=2 th34=2 th35=2;
  c3=3;    th23=(1+c3)*k3*th21*th22/(th22+c3*th21);
  c4=1;    th24=(1+c4)*k4*th21*th22/(th22+c4*th21);
  c5=1/3;  th25=(1+c5)*k5*th21*th22/(th22+c5*th21);
  th2=th21*(ray=1)+th22*(ray=2)+th23*(ray=3)+th24*(ray=4)+th25*(ray=5);
  th3=th31*(ray=1)+th32*(ray=2)+th33*(ray=3)+th34*(ray=4)+th35*(ray=5);
  t=(dd/th2)**th3; den=1+t; p=t/den;
  model y~binomial(n,p);
run;
```

The NLMIXED Procedure

Specifications

| | |
|---|---|
| Data Set | WORK.ONE |

| | |
|---|---|
| Dependent Variable | y |
| Distribution for Dependent Variable | Binomial |
| Optimization Technique | Dual Quasi-Newton |
| Integration Method | None |

Dimensions

| | |
|---|---|
| Observations Used | 20 |
| Observations Not Used | 0 |
| Total Observations | 20 |
| Parameters | 10 |

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 71.0 |
| AIC (smaller is better) | 91.0 |
| AICC (smaller is better) | 115.5 |
| BIC (smaller is better) | 101.0 |

Parameter Estimates

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| th21 | 9.9693 | 1.5017 | 20 | 6.64 | <.0001 | 0.05 | 6.8369 | 13.1017 |
| th22 | 11.5012 | 1.4087 | 20 | 8.16 | <.0001 | 0.05 | 8.5628 | 14.4397 |
| k3 | 1.4615 | 0.2055 | 20 | 7.11 | <.0001 | 0.05 | 1.0328 | 1.8903 |
| k4 | 2.5140 | 0.5527 | 20 | 4.55 | 0.0002 | 0.05 | 1.3611 | 3.6670 |
| k5 | 1.5726 | 0.2806 | 20 | 5.60 | <.0001 | 0.05 | 0.9872 | 2.1580 |
| th31 | 1.3699 | 0.2907 | 20 | 4.71 | 0.0001 | 0.05 | 0.7634 | 1.9764 |
| th32 | 1.8056 | 0.3320 | 20 | 5.44 | <.0001 | 0.05 | 1.1131 | 2.4981 |
| th33 | 2.6407 | 0.4850 | 20 | 5.44 | <.0001 | 0.05 | 1.6289 | 3.6525 |
| th34 | 1.7617 | 0.4330 | 20 | 4.07 | 0.0006 | 0.05 | 0.8584 | 2.6649 |
| th35 | 1.8748 | 0.3688 | 20 | 5.08 | <.0001 | 0.05 | 1.1056 | 2.6440 |

## Fourth NLMIXED

```
proc nlmixed data=one;
  parms th21=10 th22=15 k3=1 k4=1 k5=1 th3=2;
  c3=3;    th23=(1+c3)*k3*th21*th22/(th22+c3*th21);
  c4=1;    th24=(1+c4)*k4*th21*th22/(th22+c4*th21);
  c5=1/3;  th25=(1+c5)*k5*th21*th22/(th22+c5*th21);
  th2=th21*(ray=1)+th22*(ray=2)+th23*(ray=3)
      +th24*(ray=4)+th25*(ray=5);
  t=(dd/th2)**th3; den=1+t; p=t/den;
  model y~binomial(n,p);
  estimate 'LD50 on ray3' th23;
  estimate 'LD50 on ray4' th24;
  estimate 'LD50 on ray5' th25;
run;
```

The NLMIXED Procedure

Specifications

| | |
|---|---|
| Data Set | WORK.ONE |
| Dependent Variable | y |
| Distribution for Dependent Variable | Binomial |
| Optimization Technique | Dual Quasi-Newton |
| Integration Method | None |

**Dimensions**

| | |
|---|---|
| Observations Used | 20 |
| Observations Not Used | 0 |
| Total Observations | 20 |
| Parameters | 6 |

**Fit Statistics**

| | |
|---|---|
| -2 Log Likelihood | 76.8 |
| AIC (smaller is better) | 88.8 |
| AICC (smaller is better) | 95.2 |
| BIC (smaller is better) | 94.7 |

**Parameter Estimates**

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| th21 | 10.0510 | 1.2227 | 20 | 8.22 | <.0001 | 0.05 | 7.5005 | 12.6015 |
| th22 | 11.4984 | 1.3996 | 20 | 8.22 | <.0001 | 0.05 | 8.5790 | 14.4178 |
| k3 | 1.5678 | 0.2515 | 20 | 6.23 | <.0001 | 0.05 | 1.0433 | 2.0924 |
| k4 | 2.4615 | 0.4222 | 20 | 5.83 | <.0001 | 0.05 | 1.5808 | 3.3423 |
| k5 | 1.5739 | 0.2549 | 20 | 6.17 | <.0001 | 0.05 | 1.0422 | 2.1056 |
| th3 | 1.8192 | 0.1619 | 20 | 11.24 | <.0001 | 0.05 | 1.4816 | 2.1569 |

**Additional Estimates**

| Label | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|
| LD50 on ray3 | 17.4012 | 2.2629 | 20 | 7.69 | <.0001 | 0.05 | 12.6808 | 22.1216 |
| LD50 on ray4 | 26.4023 | 3.9178 | 20 | 6.74 | <.0001 | 0.05 | 18.2299 | 34.5747 |
| LD50 on ray5 | 16.3335 | 2.1021 | 20 | 7.77 | <.0001 | 0.05 | 11.9485 | 20.7185 |