```
(ch4-3-crab-2_HO.out)
/* Compute the slope of the log(sd) vs. log(mean) plot
   for the Box-Cox Variance Stabilizing Transformation */

*Save the mean and sd for each group in the dataset 'temp1';
 PROC MEANS NOPRINT DATA=a1;
    BY habitat;
    OUTPUT OUT=temp1 MEAN=mean STD=sd;
    RUN;
*Add the variables logsd and logmean to the dataset;
 DATA temp1;
    SET temp1;
    logsd = LOG(sd);
    logmean = LOG(mean);
    RUN;
*Plot the relationship between logsd and logmean;
 PROC PLOT VPERCENT=50 DATA=temp1;
    PLOT logsd*logmean;
    RUN;
*Find the slope of the estimated regression line;
 PROC GLM DATA=temp1;
    MODEL logsd = logmean;
    RUN;
    QUIT;
  *Note that there is no CLASS statement for GLM here;
  *This gives a regression rather than ANOVA model;

title "Log Transformed Data";

*Add the transformed variable to the dataset 'a1';
 DATA a1;
    SET a1;
    newnum = LOG(number + 1/6);
    RUN;

*Save the residuals and predicted values in the dataset 'a3';
 PROC GLM DATA=a1;
    CLASS habitat;
    MODEL  newnum = habitat;
    OUTPUT OUT=a3 PREDICTED=pred RESIDUAL=resid;
    RUN;

*Plot of residuals vs. predicted values;
 PROC PLOT DATA=a3;
    PLOT resid*pred;
    RUN;

*Compute Levene's and Brown & Forsythe's HOV tests;
 PROC GLM DATA=a1;
    CLASS habitat;
    MODEL  newnum = habitat;
    MEANS habitat / HOVTEST=LEVENE HOVTEST=BF;
    RUN;
    QUIT;
```

The GLM Procedure - LOG(sd) vs. LOG(mean) regression relationship

Dependent Variable: logsd

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	3.74311451	3.74311451	213.35	0.0001
Error	4	0.07017868	0.01754467		
Corrected Total	5	3.81329319			

R-Square	Coeff Var	Root MSE	logsd Mean
0.981596	3.546808	0.132456	3.734521

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.6559222372	0.21759635	3.01	0.0394
logmean	0.9857700086	0.06748879	14.61	0.0001
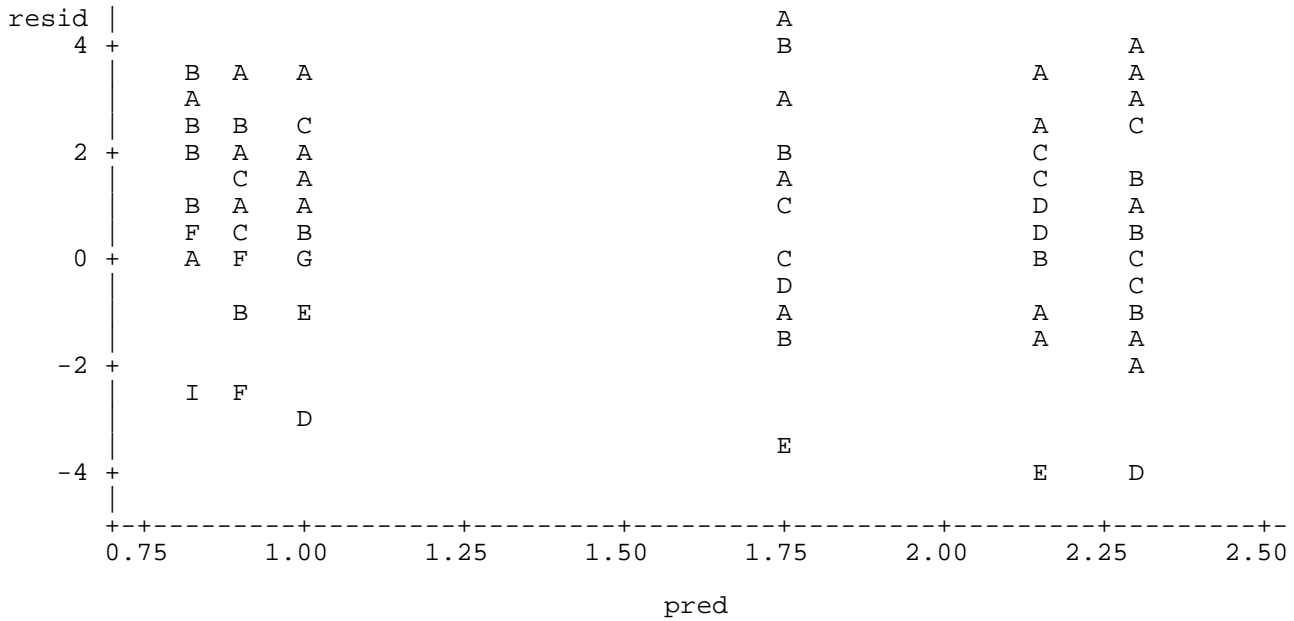
```
Log Transformed Data
--------------------

        Plot of resid*pred.   Legend: A = 1 obs, B = 2 obs, etc.

resid |                                          A
    4 +                                          B                            A
      |      B   A   A                                              A         A
      |      A                                   A                            A
      |      B   B   C                                              A         C
    2 +      B   A   A                            B                 C
      |          C   A                            A                 C         B
      |      B   A   A                            C                 D         A
      |      F   C   B                                              D         B
    0 +      A   F   G                            C                 B         C
      |                                           D                           C
      |          B   E                            A                 A         B
      |                                           B                 A         A
   -2 +                                                                       A
      |      I   F
      |              D
      |                                           E
   -4 +                                                           E         D
      |
      +-+---------+---------+---------+---------+---------+---------+---------+-
        0.75      1.00      1.25      1.50      1.75      2.00      2.25      2.50

                                      pred

Class          Levels    Values
habitat             6     1 2 3 4 5 6

Dependent Variable: newnum
                                  Sum of
Source                    DF      Squares      Mean Square    F Value    Pr > F
Model                      5     54.7253302    10.9450660       2.32     0.0460
Error                    144    678.5967209     4.7124772                ------
Corrected Total          149    733.3220511

R-Square      Coeff Var      Root MSE      newnum Mean
0.074627      145.6185       2.170824        1.490761


 Brown and Forsythe's Test for Homogeneity of newnum Variance
       ANOVA of Absolute Deviations from Group Medians

                      Sum of       Mean
Source         DF     Squares      Square    F Value    Pr > F
habitat         5     7.1934       1.4387      0.75      0.5864
Error         144     275.7        1.9149                ------


Tukey's Studentized Range (HSD) Test

Transformed Data (newnum)              Original Data (number)
-------------------------------        -------------------------------
Alpha                     0.05         Alpha                      0.05
Error D.F.                 144         Error D.F.                  144
Error Mean Square     4.71247          Error Mean Square      5170.091
Min Signif Difference  1.7735          Min Signif Difference   58.744

Grouping    Mean      N    habitat     Grouping       Mean        N    habitat

A         2.2978     25    2                  A       68.72       25    2
A         2.1612     25    1           B      A       50.64       25    3
A         1.7490     25    3           B      A       33.80       25    1
A         1.0072     25    5           B      A       12.64       25    6
A         0.9078     25    4           B      A       10.00       25    5
A         0.8215     25    6           B               9.24       25    4
```