Previous lecture

- ESS F-test
  - using sequential ANOVA table to perform test
  - connection with t-test for single coefficients in multiple regression
  - manipulate sequential ANOVA tables to produce tables for nested models, and perform tests
- Modeling interaction effects using indicator variable assumes error variance same in each group.
- CI for a linear combination of population parameters.

Variance-Covariance Matrix of Parameter Estimates

$$\begin{bmatrix} [SE(\hat{\beta}_0)]^2 & cov(\hat{\beta}_0,\hat{\beta}_1) & cov(\hat{\beta}_0,\hat{\beta}_2) & cov(\hat{\beta}_0,\hat{\beta}_3) \\ \cdots & [SE(\hat{\beta}_1)]^2 & cov(\hat{\beta}_1,\hat{\beta}_2) & cov(\hat{\beta}_1,\hat{\beta}_3) \\ \cdots & \cdots & [SE(\hat{\beta}_2)]^2 & cov(\hat{\beta}_2,\hat{\beta}_3) \\ \cdots & \cdots & \cdots & [SE(\hat{\beta}_3)]^2 \end{bmatrix}$$

Other linear combinations

- Give a 95% confidence interval for the mean percentage of trees damaged in the South when elevation=1000. Also give a prediction interval for a new observation when elevation=1000.
- Two ways to solve:
  - computer centering trick
  - write as a linear combination

Solution using computer centering

Formula is:

```
Perc.Damage~I(Elevation - 1000)+Location+Location:I(Elevation-1000)
```

Results:

```
Coefficients:
                             Value Std. Error  t value Pr(>|t|)
                 (Intercept) 20.0707   7.9916     2.5115   0.0147
          I(Elevation - 1000) -0.0172   0.0193    -0.8929   0.3755
                    Location 29.7693   8.3289     3.5742   0.0007
Location:I(Elevation - 1000)  0.1084   0.0233     4.6464   0.0000

Residual standard error: 14.61 on 60 degrees of freedom
Multiple R-Squared: 0.5436
F-statistic: 23.82 on 3 and 60 degrees of freedom, the p-value is
2.825e-010
```

$$SE(\hat{\mu}(D|E,L)) = ?$$
$$SE(Pred(D|E,L)) = ?$$

Road Map for Model Selection: Bat Data

Is there a difference in the in-flight energy expenditures of echolocating and non-echolocating bats accounting for body size?

Analysis of residuals suggests analyzing energy and mass on log scale (see previous handout for plots).

Rule out separate lines model for the three groups using test for significance of interaction effects. There are 2 interaction effects, so need ESS F-test. (See Display 10.12 and following slides.)

Consider parallel lines vs. equal lines model.

Separate Regression Lines Model for the Bat Data

options(contrasts="contr.treatment")

log(energy) ~ log(mass) + type + log(mass):type

```
Coefficients:
                Value Std. Error t value Pr(>|t|)
   (Intercept) -0.2024  1.2613    -0.1605  0.8748
     log(mass)  0.5898  0.2061     2.8611  0.0126
         type2 -1.3784  1.2952    -1.0642  0.3053
         type3  1.2681  1.2854    -0.9865  0.3406
log(mass)type2  0.2456  0.2134     1.1507  0.2691
log(mass)type3  0.2149  0.2236     0.9609  0.3529

Residual standard error: 0.1899 on 14 df
Multiple R-Squared: 0.9832
F-statistic: 163.4 on 5 and 14 degrees of freedom,
the p-value is 6.696e-012
```

Separate Regression Lines Model for the Bat Data

ANOVA TABLE

```
 Terms added sequentially

              Df Sum of Sq  Mean Sq  F Value     Pr(F)
     log(mass)  1  29.39191 29.39191 815.0382 0.0000000
         type  2   0.02957  0.01479   0.4100 0.6713410
log(mass):type  2   0.04845  0.02422   0.6718 0.5265309
     Residuals 14   0.50487  0.03606
```

Midterm practice problem: Is there sufficient evidence that the association between energy expenditure and body size differs among the 3 types of flying vertebrates? Use the information given above to form the appropriate test statistic and give its p-value. (answer on p. 274)