Stats for Clinical Trials, Math 150
Jo Hardin
Logistic Regression example:
drop-in-deviance & categorical variables

Length of Bird Nest

This example is from problem E1 in your text and includes 99 species of N. American passerine birds. Recall that the response variable is binary and represents whether there is a small opening ("closed=1") or a large opening ("closed=0") for the nest. The explanatory variable of interest was the length of the bird.

```
library(readr)
nests <- read_csv("~/Dropbox/teaching/math150/PracStatCD/Data Sets/Chapter 07/CSV Files/C7 Birdnest.csv
                  na="*")
```

Drop-in-deviance (Likelihood Ratio Test, LRT)

χ^2: The Likehood ratio test also tests whether the response is explained by the explanatory variable. We can output the deviance (= K - 2 * log-likelihood) for both the full (maximum likelihood!) and reduced (null) models.

$$\begin{aligned} G &= 2 \cdot \ln(L(MLE)) - 2 \cdot \ln(L(null)) \\ &= \text{null (restricted) deviance - residual (full model) deviance} \\ G &\sim \chi^2_\nu \quad \text{when the null hypothesis is true} \end{aligned}$$

where ν represents the difference in the number of parameters needed to estimate in the full model versus the null model.

```
summary(glm(`Closed?` ~ Length, data = nests, family="binomial"))
```

```
##
## Call:
## glm(formula = `Closed?` ~ Length, family = "binomial", data = nests)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.0581  -0.9450  -0.8039   1.3388   2.4275
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.45710    0.75301   0.607    0.544
## Length      -0.06767    0.04254  -1.591    0.112
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 119.99  on 94  degrees of freedom
## Residual deviance: 116.88  on 93  degrees of freedom
##   (4 observations deleted due to missingness)
## AIC: 120.88
##
## Number of Fisher Scoring iterations: 4
```

```r
summary(glm(`Closed?` ~ as.factor(Length), data=nests, family="binomial"))
```

```
##
## Call:
## glm(formula = `Closed?` ~ as.factor(Length), family = "binomial",
##     data = nests)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.48230  -0.75853  -0.00008   0.90052   1.97277
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)             1.957e+01  1.075e+04   0.002    0.999
## as.factor(Length)10     4.319e-07  1.317e+04   0.000    1.000
## as.factor(Length)10.5   4.296e-07  1.521e+04   0.000    1.000
## as.factor(Length)11    -1.887e+01  1.075e+04  -0.002    0.999
## as.factor(Length)12    -2.118e+01  1.075e+04  -0.002    0.998
## as.factor(Length)12.5   4.306e-07  1.521e+04   0.000    1.000
## as.factor(Length)13    -2.026e+01  1.075e+04  -0.002    0.998
## as.factor(Length)13.5  -2.066e+01  1.075e+04  -0.002    0.998
## as.factor(Length)14    -1.928e+01  1.075e+04  -0.002    0.999
## as.factor(Length)14.5  -3.913e+01  1.317e+04  -0.003    0.998
## as.factor(Length)15    -2.136e+01  1.075e+04  -0.002    0.998
## as.factor(Length)15.5  -3.913e+01  1.242e+04  -0.003    0.997
## as.factor(Length)16    -1.957e+01  1.075e+04  -0.002    0.999
## as.factor(Length)16.5  -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)17    -1.916e+01  1.075e+04  -0.002    0.999
## as.factor(Length)17.5  -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)18    -2.026e+01  1.075e+04  -0.002    0.998
## as.factor(Length)18.5  -3.913e+01  1.242e+04  -0.003    0.997
## as.factor(Length)19    -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)19.5  -1.957e+01  1.075e+04  -0.002    0.999
## as.factor(Length)20    -1.906e+01  1.075e+04  -0.002    0.999
## as.factor(Length)21    -3.913e+01  1.317e+04  -0.003    0.998
## as.factor(Length)21.5  -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)22.5  -1.957e+01  1.075e+04  -0.002    0.999
## as.factor(Length)23    -3.913e+01  1.317e+04  -0.003    0.998
## as.factor(Length)24    -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)25.5  -3.913e+01  1.317e+04  -0.003    0.998
## as.factor(Length)26.5  -3.913e+01  1.317e+04  -0.003    0.998
## as.factor(Length)29    -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)29.5  -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)30.5  -3.913e+01  1.242e+04  -0.003    0.997
## as.factor(Length)31.5  -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)48    -3.913e+01  1.521e+04  -0.003    0.998
## as.factor(Length)49.5   4.228e-07  1.521e+04   0.000    1.000
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 119.992  on 94  degrees of freedom
## Residual deviance:  73.755  on 61  degrees of freedom
##   (4 observations deleted due to missingness)
## AIC: 141.76
```

```
##
## Number of Fisher Scoring iterations: 18
```

```
summary(glm(`Closed?` ~ Length + Incubate +  Color, data=nests, family="binomial"))
```

```
##
## Call:
## glm(formula = `Closed?` ~ Length + Incubate + Color, family = "binomial",
##     data = nests)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.3712  -0.8809  -0.6958   1.3182   2.4303
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.63795    2.06443  -1.278   0.2013
## Length      -0.11412    0.05266  -2.167   0.0302 *
## Incubate     0.31392    0.17228   1.822   0.0684 .
## Color       -0.42028    0.60881  -0.690   0.4900
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 110.09  on 87  degrees of freedom
## Residual deviance: 103.33  on 84  degrees of freedom
##   (11 observations deleted due to missingness)
## AIC: 111.33
##
## Number of Fisher Scoring iterations: 4
```