Lab 3 Part 2: GLM Predictions

Your_Name_Here

Turn in via Canvas next Friday by the end of the day. Submit either a knitted pdf, html, or doc file (you can submit the .Rmd file if knitting doesn't work).

1. Read in the salamander demographics data and look at the structure. (2 pts)

```
sal <- read.csv("Data/Salamander_Demographics.csv", stringsAsFactors = FALSE)
str(sal)
```

```
## 'data.frame':	3382 obs. of  20 variables:
##  $ line  : int  1861 1115 360 2897 1432 372 231 2739 2236 543 ...
##  $ page  : int  60 36 12 92 46 12 8 87 72 17 ...
##  $ dates : chr  "4/21/09" "9/9/08" "5/31/08" "5/7/11" ...
##  $ month : int  4 9 5 5 10 5 5 10 5 6 ...
##  $ day   : int  21 9 31 7 16 31 27 24 14 5 ...
##  $ year  : int  2009 2008 2008 2011 2008 2008 2008 2009 2009 2008 ...
##  $ time  : chr  "N" "N" "N" "N" ...
##  $ plot  : chr  "5" NA "3" "7" ...
##  $ mass  : num  0.427 0.633 0.639 0.921 0.943 ...
##  $ svl   : int  33 37 42 43 45 46 47 48 NA NA ...
##  $ tl    : int  63 68 63 79 74 NA 75 89 87 NA ...
##  $ sex   : chr  NA NA NA NA ...
##  $ gravid: chr  "N" "N" "N" "N" ...
##  $ group : chr  NA NA NA NA ...
##  $ clutch: int  NA NA NA NA NA NA NA NA NA NA ...
##  $ color : chr  "R" "R" "R" "R" ...
##  $ recap : chr  NA NA NA "N" ...
##  $ mark  : chr  NA NA NA NA ...
##  $ id    : int  1371 NA 187 2154 1042 198 74 2036 1564 351 ...
##  $ damage: chr  "N" "N" "Y" "N" ...
```

2. Run a linear model to look at the effects of svl and sex on mass. Print the summary results. (2 pts)

```
lm1 <- lm(mass ~ 1 + svl + sex, data = sal)
summary(lm1)
```

```
##
## Call:
## lm(formula = mass ~ 1 + svl + sex, data = sal)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.45666 -0.06701 -0.00879  0.06072  0.67157
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -0.8538963  0.0196565 -43.441  < 2e-16 ***
## svl          0.0403860  0.0005836  69.205  < 2e-16 ***
## sexUA        0.0518362  0.0401968   1.290    0.197
## sexUI        0.1393014  0.0104140  13.376  < 2e-16 ***
## sexX         0.0479547  0.0075245   6.373  2.1e-10 ***
## sexY        -0.0006817  0.0066384  -0.103    0.918
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1117 on 3364 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.8332, Adjusted R-squared:  0.8329
## F-statistic:  3360 on 5 and 3364 DF,  p-value: < 2.2e-16
```

3. What is the expected mass of an individual of male with a snout-vent length (svl) of 40? (4 pts)

```
mass_y <- -0.8538963 + 0.0400386 * 40 + -0.00068 * 1
mass_x <- -0.8538963 + 0.0400386  * 40 + 0.047855 * 1
```

0.747 g

4. What is the expected mass of a female with the same svl? (4 pts)

A female would have a mass of 0.796 g. This is 6.5% bigger than males of the same length.

5. Run a GLM to test the effect of year, elevation, and landuse on the count (count1) of bunnies. (2 pts)

```
bunnies <- read.table("Data/hares_data.txt", header = TRUE, stringsAsFactors = FALSE)
glm1 <- glm(count1 ~ year + elevation + landuse, data = bunnies, family = "poisson")
summary(glm1)
```

```
##
## Call:
## glm(formula = count1 ~ year + elevation + landuse, family = "poisson",
##     data = bunnies)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -8.178  -4.402  -1.569   2.029  18.248
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -34.100950   3.224542 -10.575   <2e-16 ***
## year           0.018513   0.001608  11.509   <2e-16 ***
## elevation      0.001170   0.000137   8.542   <2e-16 ***
## landusegrass  -0.295820   0.018475 -16.012   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 13896  on 596  degrees of freedom
## Residual deviance: 13464  on 593  degrees of freedom
##   (355 observations deleted due to missingness)
## AIC: 16288
##
## Number of Fisher Scoring iterations: 5
```

6. What is the expected number of bunnies on at low elevation in a grassland in 2005? What about at high elevation? (6 pts)

```r
min(bunnies$elevation, na.rm = T)
max(bunnies$elevation, na.rm = T)
exp(-34.1 + 0.018513 * 2005 + 0.001170 * 350 + -0.295820 * 1)
exp(-34.1 + 0.018513 * 2005 + 0.001170 * 600 + -0.295820 * 1)
```

The elevation in the data range from 350 to 622 meters. So the expected number of hares at a low elevation of 350 m in a grassland in 2005 would be 23. We use the exponential because of the log link within a Poisson GLM. The expected number in the same year and habitat (grassland) would be 31 at a high elevation of 600 m