Test 2

Trevor VanVeldhuisen

November 01, 2021

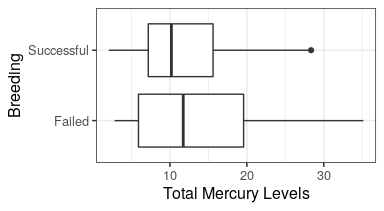
birds <- read\_csv('https://sldr.netlify.app/data/seabird-mercury.csv')   
 glimpse(birds)

## Rows: 229  
## Columns: 8  
## $ bird\_ID <dbl> 1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6,…  
## $ total\_mercury <dbl> 15.015, 15.641, 6.867, 6.949, 12.497, 12.501, 8.275, …  
## $ foraging\_habitat <dbl> -20.07, -20.01, -19.63, -20.33, -20.22, -19.99, -18.6…  
## $ trophic\_level <dbl> 13.82, 13.76, 11.50, 10.25, 13.56, 13.91, 13.27, 13.5…  
## $ age <dbl> 35, 35, 35, 35, 35, 35, 32, 32, 32, 32, 32, 32, 32, 3…  
## $ sex <chr> "M", "M", "M", "M", "M", "M", "F", "F", "F", "F", "F"…  
## $ breeding\_history <chr> "Failed\_Not\_seen", "Failed\_Not\_seen", "Failed\_Not\_see…  
## $ breeding <chr> "Failed", "Failed", "Failed", "Successful", "Successf…

The specific research question I am going to answer is if mercury levels have an effect on the success of breeding in grey-headed albatrosses near Antarctica. Response Variable: breeding success Predictors: total\_mercury, breeding\_history, age, foraging\_habitat, trophic\_level, sex Total Mercury is the main predictor variable that we want to see if it has an association with breeding success. I included the rest of the variables to see if there are other factors that might also be associated with breeding success.

birds <- birds %>%  
 mutate(breeding = factor(breeding),  
 breeding\_history = factor(breeding\_history))

gf\_boxplot( breeding ~ total\_mercury, data = birds) %>%  
 gf\_labs( y = 'Breeding',  
 x = 'Total Mercury Levels')



From this graph we learn that successful breeding tends to happen when here are lower levels of mercury. The median and the upper quartile of failed breeding are both greater than successful breeding.

birds.logr <- glm(factor(breeding) ~ total\_mercury + foraging\_habitat + trophic\_level + age + sex + breeding\_history,   
 data = birds,  
 family = binomial(link='logit'))

summary(birds.logr)

##   
## Call:  
## glm(formula = factor(breeding) ~ total\_mercury + foraging\_habitat +   
## trophic\_level + age + sex + breeding\_history, family = binomial(link = "logit"),   
## data = birds)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -1.7333 -0.7521 -0.5125 0.9113 2.3246   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -10.84867 1070.95927 -0.010 0.9919   
## total\_mercury -0.05375 0.02543 -2.114 0.0346 \*  
## foraging\_habitat 0.28760 0.15317 1.878 0.0604 .  
## trophic\_level 0.09420 0.15007 0.628 0.5302   
## age -0.07102 0.03269 -2.173 0.0298 \*  
## sexM 0.30153 0.35815 0.842 0.3998   
## breeding\_historyFailed\_Not\_seen 16.22425 1070.95055 0.015 0.9879   
## breeding\_historyNot\_seen\_Failed -0.22808 1651.65940 0.000 0.9999   
## breeding\_historyNot\_seen\_Not\_seen 16.30239 1070.95059 0.015 0.9879   
## breeding\_historySuccessful\_Not\_seen 17.97757 1070.95057 0.017 0.9866   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 278.58 on 228 degrees of freedom  
## Residual deviance: 225.77 on 219 degrees of freedom  
## AIC: 245.77  
##   
## Number of Fisher Scoring iterations: 16

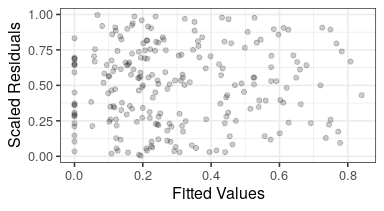
logit(pi) = -10.84 -0.05375total\_mercury + 0.28760foraging\_habitat + 0.09420trophic\_level - 0.07102age + 0.030153sexM + 16.22425breeding\_historyFailed\_Not\_seen - 0.22808breeding\_historyNot\_seen\_Failed + 16.30239breeding\_historyNot\_seen\_Not\_seen + 1797757breeding\_historySuccessful\_Not\_seen + E yi ~ Binom(n, pi)

I chose binary regression because I have a binary response variable: breeding. The logit funcation was also used because the outcome is binary.

require(DHARMa)

birds\_sim <- (simulateResiduals(birds.logr))

gf\_point(birds\_sim$scaledResiduals ~ fitted(birds.logr),  
 alpha = 0.2) %>%  
 gf\_labs(x = 'Fitted Values',  
 y = 'Scaled Residuals')

 Through this scaled Residuals vs Fitted Values plot we see that there are no trends and the scaled residuals have uniform distribution showing it passes the mean-variance test.

car::Anova(birds.logr)

## Analysis of Deviance Table (Type II tests)  
##   
## Response: factor(breeding)  
## LR Chisq Df Pr(>Chisq)   
## total\_mercury 4.829 1 0.02798 \*   
## foraging\_habitat 3.612 1 0.05737 .   
## trophic\_level 0.398 1 0.52829   
## age 4.982 1 0.02562 \*   
## sex 0.713 1 0.39833   
## breeding\_history 40.641 4 3.19e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

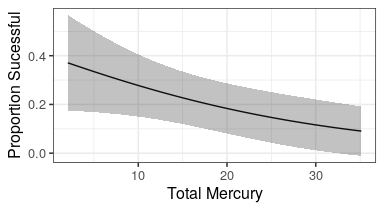
The null hypothesis is that there is not association between total mercury and breeding success. Through testing the null hypothesis, we have evidence to reject the null hypothesis since there is association between total mercury and breeding success as well as age and breeding success.

require(s245)

get\_fixed(birds.logr)

## factor.breeding. total\_mercury foraging\_habitat trophic\_level age sex  
## 1 Failed 11.057 -20.08 13.19 22 M  
## breeding\_history  
## 1 Failed\_Not\_seen

pred\_plot(birds.logr, 'total\_mercury', data = birds) %>%  
 gf\_labs( y = 'Proportion Sucessful',  
 x = 'Total Mercury')



This prediction plot shows a negative correlation between successful breeding and total mercury levels. The more mercury content in a grey-headed albatross, the less likely they are able to successfully breed.

Looking at all of the data presented above, it seems that mercury content in grey-headed albatrosses does have a negative effect on the breeding success. First, looking at the boxplot, we can see the median mercury level is higher in albatrosses who were not successful in breeding. We can also observe the upper quartile extending beyond that of those who were successful at breeding. Second, looking at the p-value of mercury (-0.05375), we can assume that total\_mercury is a significant predictor. One other significant predictor was age (p = -0.07102). The rest of the predictors I chose: foraging\_habitat, trophic\_level, sex, and breeding\_history, were not as strong and had higher p-values suggesting they are not significant predictors. Third, the model passed the mean-variance test, as there were no patterns or trends, and the scaled residuals showed a uniform distribution. Fourth, running a hypothesis test, we concluded to reject the null hypothesis since there is evidence of an association between total\_mercury and breeding success. We also found that age and breeding success also have an association through hypothesis testing. Fifth, looking at the prediction plot we can see a negative correlation between total\_mercury and the proportion of successful breeding. lastly, looking at the odds ratio, the Beta for Mercury is -0.05375. The exponentiation of e^-0.05375 = 0.94766 is the odds ratio for a one unit change in mercury. The odds of breeding are 0.94 less for every one unit change in mercury showing a negative association between breeding success and total\_mercury.