Test 3

Trevor VanVeldhuisen

December 05, 2021

1. DATA

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…

birds <- birds %>%  
 arrange(bird\_ID)

The specific research question I am going to answer is if mercury levels are associated with the success of breeding in grey-headed albatrosses near Antarctica. I am going to answer this question by fitting a binary regression model since the outcome of interest here is binary. The model plan includes the response variable: breeding success and the following set of 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. However, I included the rest of the variables that describe features of the albatrosses to see if these other factors might also be associated with breeding, as total mercury may not be the only factor and others may provide insight as to the impact of breeding success. After model fitting, the breeding history predictor provided peculiar estimates. Since it was not part of the primary question, I decided to remove breeding history from the model.

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

1. RATIONALE The random effect chosen is Bird ID. This is because the data are dependent on the specific bird, so to account for the dependence we use a random effect based on each bird.
2. FITTING With Random Effect

birds.logrRE <- glmmTMB(factor(breeding) ~ total\_mercury + foraging\_habitat + trophic\_level + age + sex + (1|bird\_ID),   
 data = birds,  
 family = binomial(link='logit'))

Without Random Effect

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

SUMMARY

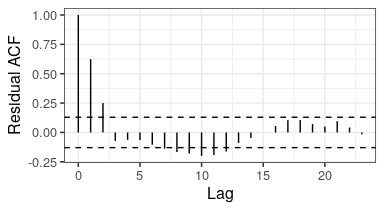
summary(birds.logrRE)

## Family: binomial ( logit )  
## Formula:   
## factor(breeding) ~ total\_mercury + foraging\_habitat + trophic\_level +   
## age + sex + (1 | bird\_ID)  
## Data: birds  
##   
## AIC BIC logLik deviance df.resid   
## 86.9 110.9 -36.5 72.9 222   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## bird\_ID (Intercept) 3273 57.21   
## Number of obs: 229, groups: bird\_ID, 78  
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -8.59602 28.58400 -0.301 0.764  
## total\_mercury -0.04710 0.16884 -0.279 0.780  
## foraging\_habitat 0.18924 1.02763 0.184 0.854  
## trophic\_level 0.07517 0.96129 0.078 0.938  
## age -0.05597 0.26262 -0.213 0.831  
## sexM 0.88710 2.97191 0.298 0.765

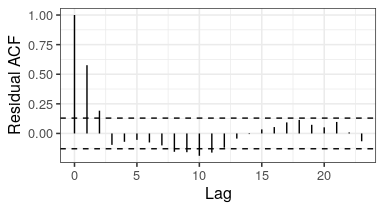
Model Equation: -8.596 - 0.047total\_mercury + 0.189foraging\_habitat + 0.075trophic\_level - 0.056age + 0.887sexM + Erfbird\_ID + Eresid Variance Estimate is 3273 and is the estimate of variance explained by having the same birds for multiple observations. We have 78 unique birds that contribute 229 observations so this variance estimate accounts for having multiple birds for multiple observations.

ASSESSMENT Without Random Effect

gf\_acf(~birds.logr)

 With Random Effect

gf\_acf(~birds.logrRE)

 Looking at these two ACF tests, the conditions are not met as the there are some residuals that exceed beyond the confidence bounds, and also seems to have a pattern to it as well. The random effects model did not change much from the model without the random effect.

PREDICTIONS For this model, it would be better to make population type predictions because. We want to use these results to generalize the data to all grey-headed albatross’ in Antarctica which is why we include the random effect. For this data set specifically, the random effect did not change much in the dependence of the data because there were many birds being observed for the size of the data set so we already could have made predictions about the population average.

EXTRA CREDIT The pros of having these tests is seeing how we can apply what we learned in class without time restrictions. Especially when writing code and analyzing, it takes more time than one would think. It was nice to be able to use resources as the coding on top of the material was a lot to take in at once. Some cons of these style tests are that you feel like you do not need to study as much for this class, and most things are in the notes but not in memory. It is more about applying what we learned rather than memorizing facts (which I respond to better but still could be a con for others).