A Tutorial for the package CAXA (conspecific attraction experiment analysis)

Sean E. MacDonald

December 14, 2017

1 Introduction

The package contains three functions that help clean and analyze a specific dataset pertaining to a series of conspecific attraction experiments. These series of 80 experiments were designed to investigate the efficacy of attracting fruit-eating birds to fruiting plants. More specifically, these experiments tease apart whether or not birds in this system utilize social information (i.e. vocalizations) when making foraging decisions. And if so, do they rely on conspecific cues, heterospecific cues, or both? Do members within the frugivory dietarty guild utilize eavesdropping when exploiting resources? These are the questions this small study aimed to answer. Birds within 10 meters of the focal were recorded. The three functions are as follows:

The Remove.WRSH.ZEDO function removes experiments containing trials of the track species White-rumped shama (WRSH) and Zebra dove (ZEDO) due to low sampling effort. Low samples were taken due to diet analysis determining that these species are primarily insectivorous or don't respond to conspecific vocalizations.

The Keep.15.Tracks function keeps experiments containing trials equal to 15 minutes in order to standardize methodology and low sampling efforts with track lengths not equal to 15 minutes. Preliminary experiments utilized various track lengths until a standardized methodology was implemented.

The GLMM.Bin function applys a generalized linear mixed model (GLMM) with a binary (Bin) distribution to the data in order to see what bird species significantly respond to track species. Applying this type of model is necessary as the data is not normally distributed (generalized). Additionally, trials were not independent of one another meaning track order must be categorized as a random effect. Therefore, this model will contain both fixed and random effects, hence mixed.

2 Dataset

The dataset is a product of a series of 80 experiments conducted on the Island of Oahu, Hawaii, USA from June 2016 to July 2017. These experiments were designed to investigate the efficacy of attracting fruit-eating birds to fruiting plants. More specifically, on the island of Oahu all fruit eating birds are extinct, but several exotic birds have established populations across the islands. The study investigated whether these introduced birds could be enticed to consume fruit from exotic, native, and endangered plants. Sampling took place across 26 plant species with each experiment on a different individual. Birds were only recorded within 10 meters of focal plant due to low visibility in the cluttered forest. Bird breeding season ranges from approximately March - August in Hawaii. Many more data were collected that are not present in this dataset. Initial experiments utilized track species White-rumped shama (wrsh) and Zebra dove (zedo), but will be removed due to low sampling effort. Low samples were taken due to diet analysis determining that these species are primarily insectivorous or don't respond to conspecific vocalizations. Additionally, experiments with track lengths not equal to 15 minutes will be removed due to low sampling effort. Preliminary experiments utilized various track lengths until a standardized methodology was implemented. Control periods are divided into four, 15-min trials for proper comparision when applying a generalized linear mixed model (GLMM). Applying this type of model is necessary as the data is not normally distributed and that trials were not independent of one another. More specifically, the broadcasting of the first playback influences each subsequent trial and their respective birds' behavioral response strength. As such, track order must be accounted for as a random effect hence the mixed model approach. Fixed effects that were accounted for and of particular interest include breeding season, plant origin, and track species. The latter being the more important in order to determine if birds utilize social information only from their own species' vocalizations or others within their dietary guild (i.e. frugivores).

The dataframe contains 80 experiments across 624 rows and 12 columns. Each experiment consists of a control (con) and treatment (treat) period specified under track type. Control and treatment periods have specified lengths found under track length. Treatment trials have certain bird species calls associated under the track species column as either: Japanese white-eye (jawe), Japanese white-eye and Red-billed leiothrix (jawe.rble), Red-billed leiothrix (rble), Red-vented bulbul (rvbu), and Red-whiskered bulbul (rwbu). Control periods are denoted as 'con' under track species. Each trial is associated with an order found under track order. Whether or not an experiment was conducted during birds' breeding season (0 = no, 1 = yes) is denoted under the breeding.season column. Additionally, data regarding the plant species origin is denoted in the native plant column (0 = no, 1 = yes). Lastly, behavioral response to track species playback is denoted by each focal bird species (0 = no response, 1 = attracted): Japanese white-eye (jawe), Red-billed leiothrix (rble), Red-vented bulbul (rvbu), Red-whiskered bulbul (rwbu), and all non-fruit eating species (other).

3 Example

```
> ## Load the required package and dataset
> library(CAXA)
> data("cax_data")
> ## Investigate track.spp to see how many experiments utilized
> ## each species; notice how few used wrsh and zedo; let's
> ## remove those few experiments
> table(cax_data$track.spp)

  control      jawe jawe.rble      rble      rvbu      rwbu      wrsh      zedo
      320        65        34        48        81        69         6         1

> ## Remove all experiments with track species containing wrsh or
> ## zedo. The newly cleaned dataframe (cax_data_1) should contain
> ## 573 observations now
> cax_data_1 <- Remove.WRSH.ZEDO(cax_data)
> print(cax_data_1)

# A tibble: 573 x 12
# Groups:   experiment [73]
   experiment track.order track.spp track.type track.length breeding.season
       <fctr>      <fctr>     <chr>      <chr>        <int>           <int>
 1          1           1   control        con           15               1
 2          1           2   control        con           15               1
 3          1           3   control        con           15               1
 4          1           4   control        con           15               1
 5          1           5      rvbu      treat           60               1
 6          2           1   control        con           15               1
 7          2           2   control        con           15               1
 8          2           3   control        con           15               1
 9          2           4   control        con           15               1
10          2           5      jawe      treat           15               1
# ... with 563 more rows, and 6 more variables: native.plant <int>, jawe <int>,
#   rble <int>, rvbu <int>, rwbu <int>, other <int>

> ## Investigate track.length to see how many experiments utilized
> ## different lengths; notice how most are 15 minutes long;
> ## let's keep only those experiments
> table(cax_data$track.length)

  5  10  15  20  25  30  45  60
  2  18 572  23   1   3   3   2

> ## Utilizing the cleaned dataframe (cax_data_1) from
> ## Remove.WRSH.ZEDO, this function will now Keep only experiments
> ## with all trials of track lengths equal to 15 minutes. The newly
```

```
> ## cleaned dataframe (cax_data_2) should contain 448 observations now
> cax_data_2 <- Keep.15.Tracks(cax_data_1)
> print(cax_data_2)

# A tibble: 448 x 12
# Groups:   experiment [56]
   experiment track.order track.spp track.type track.length breeding.season
       <fctr>      <fctr>     <chr>      <chr>        <int>           <int>
 1         13           1   control        con           15               0
 2         13           2   control        con           15               0
 3         13           3   control        con           15               0
 4         13           4   control        con           15               0
 5         13           5      rvbu      treat           15               0
 6         13           6 jawe.rble      treat           15               0
 7         13           7      rwbu      treat           15               0
 8         13           8      rble      treat           15               0
 9         15           1   control        con           15               0
10         15           2   control        con           15               0
# ... with 438 more rows, and 6 more variables: native.plant <int>, jawe <int>,
#   rble <int>, rvbu <int>, rwbu <int>, other <int>

> ## Utilizing the latest cleaned dataframe (cax_data_2), this function
> ## applys a generalized linear mixed model to Japanese white-eye (jawe)
> model1 <- GLMM.Bin(cax_data_2$jawe)
> ## The output shows that Japanese white-eye (jawe), when present in the
> ## immediate area are significantly attracted to playbacks of jawe,
> ## jawe.rble, rble, and rvbu in comparison to the intercept (i.e. control).
> ## The breeding season or plant origin did not signigicantly influence
> ## strength of response. Interpretation: this species may use both
> ## conspecific and heterospecific vocalizations when making foraging
> ## decisions.
> summary(model1)

Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: bird.spp ~ track.spp + breeding.season + native.plant + (1 |
    track.order)
   Data: cax_data_2

     AIC      BIC   logLik deviance df.resid
     498      535     -240      480      439

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.3400 -0.5263 -0.4832  0.4631  2.1722
```

```
Random effects:
 Groups      Name        Variance Std.Dev.
 track.order (Intercept) 0.01191  0.1091
Number of obs: 448, groups:  track.order, 8

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         -1.2871     0.2805  -4.589 4.46e-06 ***
track.sppjawe        2.9869     0.4133   7.226 4.96e-13 ***
track.sppjawe.rble   3.0937     0.6680   4.632 3.63e-06 ***
track.spprble        1.2887     0.3783   3.407 0.000658 ***
track.spprvbu        1.1971     0.3240   3.695 0.000220 ***
track.spprwbu        0.2429     0.3705   0.656 0.511997
breeding.season     -0.1335     0.2302  -0.580 0.562028
native.plant        -0.0773     0.2610  -0.296 0.767146
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) trck.sppj trck.. trck.spprb trck.spprv trck.spprw brdng.
track.sppjw -0.212
trck.sppjw. -0.275  0.106
trck.spprbl -0.257  0.217     0.143
trck.spprvb -0.334  0.261     0.147  0.271
trck.spprwb -0.316  0.217     0.155  0.260      0.278
breedng.ssn -0.353 -0.063     0.137 -0.074     -0.016      0.012
native.plnt -0.655 -0.044     0.062 -0.027     -0.001      0.005     -0.084

> ## Utilizing the latest cleaned dataframe (cax_data_2), this function
> ## applys a generalized linear mixed model to Red-billed leiothrix (rble)
> model2 <- GLMM.Bin(cax_data_2$rble)
> ## The output shows that Red-billed leiothrix (rble), when present in the
> ## immediate area are significantly attracted to playbacks of jawe.rble
> ## and rble in comparison to the intercept (i.e. control). The breeding
> ## season did not signigicantly influence strength of response, but
> ## plant origin did. Interpretation: this species may use conspecific
> ## vocalizations when making foraging decisions and prefer exotic plants.
> summary(model2)

Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: bird.spp ~ track.spp + breeding.season + native.plant + (1 |
    track.order)
   Data: cax_data_2
```

```
     AIC      BIC   logLik deviance df.resid
   419.2    456.1   -200.6    401.2      439

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.1317 -0.4588 -0.4075 -0.3375  3.4634

Random effects:
 Groups      Name        Variance Std.Dev.
 track.order (Intercept) 1.96e-12 1.4e-06
Number of obs: 448, groups:  track.order, 8

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         -1.5663     0.2982  -5.253  1.5e-07 ***
track.sppjawe        0.2372     0.4172   0.569 0.569661
track.sppjawe.rble   1.2277     0.5276   2.327 0.019956 *
track.spprble        1.4369     0.3829   3.753 0.000175 ***
track.spprvbu        0.4503     0.3791   1.188 0.234936
track.spprwbu       -0.3120     0.4754  -0.656 0.511600
breeding.season      0.3769     0.2649   1.423 0.154779
native.plant        -0.6061     0.2761  -2.195 0.028156 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) trck.sppj trck.. trck.spprb trck.spprv trck.spprw brdng.
track.sppjw -0.246
trck.sppjw. -0.389  0.153
trck.spprbl -0.251  0.237     0.164
trck.spprvb -0.327  0.233     0.186  0.254
trck.spprwb -0.266  0.186     0.148  0.202      0.205
breedng.ssn -0.448 -0.046     0.225 -0.050      0.010      0.001
native.plnt -0.572 -0.043     0.064 -0.076     -0.007      0.009     -0.099

> ## Utilizing the latest cleaned dataframe (cax_data_2), this function
> ## applys a generalized linear mixed model to Red-vented bulbul (rvbu)
> model3 <- GLMM.Bin(cax_data_2$rvbu)
> ## The output shows that Red-vented bulbul (rvbu), when present in the
> ## immediate area are significantly attracted to playbacks of jawe.rble
> ## and rvbu in comparison to the intercept (i.e. control). Plant origin
> ## did not signigicantly influence strength of response, but breeding
> ## season did. Interpretation: this species may use both
> ## conspecific and heterospecific vocalizations when making foraging
> ## decisions and be more gregarious during the breeding season.
> summary(model3)
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: bird.spp ~ track.spp + breeding.season + native.plant + (1 |
    track.order)
   Data: cax_data_2

     AIC      BIC   logLik deviance df.resid
   275.3    312.3   -128.7    257.3      439

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.1286 -0.3220 -0.2358 -0.1541  7.1080

Random effects:
 Groups      Name        Variance Std.Dev.
 track.order (Intercept) 0        0
Number of obs: 448, groups:  track.order, 8

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)         -3.5200     0.4824  -7.296 2.96e-13 ***
track.sppjawe        0.7751     0.5327   1.455 0.145637
track.sppjawe.rble   1.7244     0.7395   2.332 0.019701 *
track.spprble        0.5898     0.6125   0.963 0.335554
track.spprvbu        2.7290     0.4149   6.577 4.81e-11 ***
track.spprwbu       -0.4025     0.7837  -0.514 0.607536
breeding.season      1.2537     0.3745   3.348 0.000814 ***
native.plant        -0.2208     0.3759  -0.587 0.556892
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) trck.sppj trck.. trck.spprb trck.spprv trck.spprw brdng.
track.sppjw -0.318
trck.sppjw. -0.469  0.217
trck.spprbl -0.270  0.276     0.184
trck.spprvb -0.542  0.401     0.343  0.345
trck.spprwb -0.238  0.214     0.155  0.186      0.275
breedng.ssn -0.537 -0.022     0.284 -0.040      0.182     -0.001
native.plnt -0.515 -0.029     0.087 -0.016     -0.022      0.004     -0.100

> ## Utilizing the latest cleaned dataframe (cax_data_2), this function
> ## applys a generalized linear mixed model to Red-whiskered bulbul (rwbu)
> model4 <- GLMM.Bin(cax_data_2$rwbu)
> ## The output shows that Red-whiskered bulbul (rwbu), when present in the
```

```
> ## immediate area are significantly attracted to playbacks of rwbu in
> ## comparison to the intercept (i.e. control). Plant origin did not
> ## signigicantly influence strength of response, but breeding season did.
> ## Interpretation: this species may use conspecific vocalizations when
> ## making foraging decisions and be more grergarious during the breeding
> ## season.
> summary(model4)

Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: bird.spp ~ track.spp + breeding.season + native.plant + (1 |
    track.order)
   Data: cax_data_2

     AIC      BIC   logLik deviance df.resid
   395.6    432.5   -188.8    377.6      439

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.0237 -0.4483 -0.3578 -0.2854  3.6588

Random effects:
 Groups      Name        Variance Std.Dev.
 track.order (Intercept) 0.04627  0.2151
Number of obs: 448, groups:  track.order, 8

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)       -2.44776    0.37482  -6.531 6.55e-11 ***
track.sppjawe      0.42222    0.44633   0.946   0.3442
track.sppjawe.rble 0.11857    0.80737   0.147   0.8832
track.spprble     -0.27572    0.59465  -0.464   0.6429
track.spprvbu      0.37237    0.43443   0.857   0.3914
track.spprwbu      1.53658    0.38772   3.963 7.40e-05 ***
breeding.season    0.72045    0.27932   2.579   0.0099 **
native.plant       0.08278    0.31396   0.264   0.7920
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) trck.sppj trck.. trck.spprb trck.spprv trck.spprw brdng.
track.sppjw -0.263
trck.sppjw. -0.306  0.161
trck.spprbl -0.218  0.227     0.129
trck.spprvb -0.334  0.342     0.185  0.246
```

```
trck.spprwb -0.430  0.342     0.233  0.292      0.379
breedng.ssn -0.441 -0.050     0.166 -0.043      0.012      0.084
native.plnt -0.610 -0.028     0.067 -0.014      0.007      0.013     -0.077
```

```
> ## Utilizing the latest cleaned dataframe (cax_data_2), this function
> ## applys a generalized linear mixed model to non-frugivorous birds (other)
> model5 <- GLMM.Bin(cax_data_2$other)
> ## The output shows that non-frugivorous birds (other), when present in
> ## the immediate area ara significantly attracted to playbacks of
> ## jawe.rble in comparison to the intercept (i.e. control). Plant origin
> ## did not signigicantly influence strength of response, but breeding
> ## season did. Interpretation: these species may respond to alarm
> ## vocalizations of heterospecifics.
> summary(model5)
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: bird.spp ~ track.spp + breeding.season + native.plant + (1 |
    track.order)
   Data: cax_data_2

     AIC      BIC   logLik deviance df.resid
   526.4    563.3   -254.2    508.4      439

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.0112 -0.6488 -0.5599  1.2545  2.6890

Random effects:
 Groups      Name        Variance Std.Dev.
 track.order (Intercept) 4e-14    2e-07
Number of obs: 448, groups:  track.order, 8

Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)        -1.64406    0.28985  -5.672 1.41e-08 ***
track.sppjawe      -0.09785    0.35737  -0.274   0.7842
track.sppjawe.rble  1.09847    0.49369   2.225   0.0261 *
track.spprble      -0.06963    0.40052  -0.174   0.8620
track.spprvbu       0.34214    0.32147   1.064   0.2872
track.spprwbu      -0.33424    0.37231  -0.898   0.3693
breeding.season     0.28066    0.22403   1.253   0.2103
native.plant        0.56783    0.27318   2.079   0.0377 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Correlation of Fixed Effects:
            (Intr) trck.sppj trck.. trck.spprb trck.spprv trck.spprw brdng.
track.sppjw -0.185
trck.sppjw. -0.341  0.121
trck.spprbl -0.161  0.169     0.104
trck.spprvb -0.268  0.205     0.152  0.183
trck.spprwb -0.216  0.178     0.128  0.158      0.197
breedng.ssn -0.381 -0.043     0.198 -0.063      0.006      0.000
native.plnt -0.731 -0.033     0.119 -0.020      0.016     -0.004     -0.066
```

4 Contact Information

Please contact author via email: erroll4@illinois.edu