

# glms for proportion abscised

E.E. Jackson

06 May, 2020

```
rm(list = ls())
require(tidyverse)
require(lme4)
library(knitr)
library(report)
require(MuMIn)
require(ggpubr)
require(MASS)
theme_set(theme_pubr(base_size = 15))
opts_chunk$set(tidy.opts=list(width.cutoff=50),tidy=TRUE)
```

## Read dataset

```
fruitDat <- read.csv("../output/tables/cleanData.csv")
dm <- read.csv("../output/tables/dispersalMode.csv")
fruitDatFull <- read.csv("../output/tables/fullCleanData.csv")

fruitDat <- fruitDat %>% mutate(seedpred_pres = factor(ifelse(seedpred_pres ==
  0, "No", "Yes"))) %>% left_join(dplyr::select(dm,
  dispersal_mode = bio, sp), by = "sp")

## Warning: Column `sp` joining factors with different levels, coercing to
## character vector

fruitDatFull <- fruitDatFull %>% mutate(seedpred_pres = factor(ifelse(seedpred_pres ==
  0, "No", "Yes"))) %>% left_join(dplyr::select(dm,
  dispersal_mode = bio, sp), by = "sp")

## Warning: Column `sp` joining factors with different levels, coercing to
## character vector
```

## mixed effect models

```
# model <-
# lmer(cbind(abscised_seeds,viable_seeds)~seedpred_pres
# + (1/sp),quasibinomial,data=fruitDat)

# Error in if (REML) p else OL : argument is not
# interpretable as logical

# doesn't like lmer, change to glmer
```

```

# model <-
# glmer(cbind(abscised_seeds,viable_seeds)~seedpred_pres
# + (1|sp),quasibinomial,data=fruitDat)

# Error in lme4::glFormula(formula =
# cbind(abscised_seeds, viable_seeds) ~ : 'quasi'
# families cannot be used in glmer

# doesn't like quasibinomial, change to binomial

modell1 <- glmer(cbind(abscised_seeds, viable_seeds) ~
  seedpred_pres + (1 | sp), binomial, data = fruitDat)

## Warning in eval(family$initialize, rho): non-integer counts in a binomial glm!

summary(modell1)

## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: cbind(abscised_seeds, viable_seeds) ~ seedpred_pres + (1 | sp)
## Data: fruitDat
##
##      AIC      BIC    logLik deviance df.resid
##  2921.1   2931.3  -1457.5   2915.1     218
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.62538 -0.00484  0.00045  0.00488  0.42557
##
## Random effects:
##  Groups Name      Variance Std.Dev.
##  sp      (Intercept) 3.015   1.736
## Number of obs: 221, groups: sp, 137
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.2300    0.2780  -4.424  9.7e-06 ***
## seedpred_presYes  0.8565    0.3294   2.600  0.00931 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## sdprd_prsYs -0.844

# cbind does work! lets see if there's a difference
# in the summary output

modell2 <- glmer(proportion_abscised ~ seedpred_pres +
  (1 | sp), binomial, data = fruitDat, weights = total_seeds)

## Warning in eval(family$initialize, rho): non-integer #successes in a binomial
## glm!

```

```
summary(model2)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: proportion_abcised ~ seedpred_pres + (1 | sp)
## Data: fruitDat
## Weights: total_seeds
##
##      AIC      BIC    logLik deviance df.resid
##  2898.4   2908.6  -1446.2   2892.4     218
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.74362 -0.00561  0.00131  0.00682  0.06455
##
## Random effects:
## Groups Name      Variance Std.Dev.
## sp      (Intercept) 2.477    1.574
## Number of obs: 221, groups: sp, 137
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.5211     0.2523  -6.028 1.66e-09 ***
## seedpred_presYes  0.8535     0.2988   2.856 0.00429 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## sdprd_prsYs -0.844
```

```
# try year as random effect
```

```
model3 <- glmer(cbind(abcised_seeds, viable_seeds) ~
  seedpred_pres + (year | sp), binomial, data = fruitDatFull)
```

```
## Warning in eval(family$initialize, rho): non-integer counts in a binomial glm!
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.263596 (tol = 0.002, component 1)
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidentifiable:
## - Rescale variables?;Model is nearly unidentifiable: large eigenvalue ratio
## - Rescale variables?
```

```
r.squaredGLMM(model1) # calculate R squared for the cbind model
```

```
## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help page.
## Warning: The null model is correct only if all variables used by the original
## model remain unchanged.
## Warning in eval(family$initialize, rho): non-integer counts in a binomial glm!
##
##              R2m      R2c
## theoretical 0.04701081 0.9999792
## delta      0.04701053 0.9999732
```

```

# marginal R squared (fixed effects alone) is 0.02
# conditional R squared is 0.44

# I found out that although lme4 won't let us fit
# quasi-likelihoods, MASS will. I don't know if
# this is better or not but it allows year as a
# random effect
model4 <- MASS::glmmPQL(cbind(abscised_seeds, viable_seeds) ~
  seedpred_pres, random = ~1 | sp, family = quasibinomial,
  data = fruitDat)

## iteration 1
## iteration 2
## iteration 3
## iteration 4
## iteration 5
## iteration 6
## iteration 7
## iteration 8
## iteration 9
## iteration 10
summary(model4)

## Linear mixed-effects model fit by maximum likelihood
## Data: fruitDat
## AIC BIC logLik
## NA NA NA
##
## Random effects:
## Formula: ~1 | sp
## (Intercept) Residual
## StdDev: 2.876586 1.393261e-14
##
## Variance function:
## Structure: fixed weights
## Formula: ~invwt
## Fixed effects: cbind(abscised_seeds, viable_seeds) ~ seedpred_pres
## Value Std.Error DF t-value p-value
## (Intercept) -0.928796 0.3476311 135 -2.671786 0.0085
## seedpred_presYes 2.270076 0.3619484 135 6.271823 0.0000
## Correlation:
## (Intr)
## seedpred_presYes -0.96
##
## Standardized Within-Group Residuals:
## Min Q1 Med Q3 Max
## -4.40915821 -0.56182492 -0.03174776 0.00000000 4.82036172
##
## Number of Observations: 221

```

```

## Number of Groups: 137
model5 <- MASS::glmmPQL(cbind(abscised_seeds, viable_seeds) ~
  seedpred_pres, random = ~1 | year/sp, family = quasibinomial,
  data = fruitDatFull)

## iteration 1
## iteration 2
## iteration 3
## iteration 4
## iteration 5
## iteration 6
## iteration 7
## iteration 8
## iteration 9
## iteration 10
summary(model5)

## Linear mixed-effects model fit by maximum likelihood
## Data: fruitDatFull
## AIC BIC logLik
## NA NA NA
##
## Random effects:
## Formula: ~1 | year
## (Intercept)
## StdDev: 8.287156e-16
##
## Formula: ~1 | sp %in% year
## (Intercept) Residual
## StdDev: 6.728452 2.264051e-15
##
## Variance function:
## Structure: fixed weights
## Formula: ~invwt
## Fixed effects: cbind(abscised_seeds, viable_seeds) ~ seedpred_pres
## Value Std.Error DF t-value p-value
## (Intercept) -0.4382558 0.06133601 3750 -7.145164 0
## seedpred_presYes 0.7706134 0.06416244 3750 12.010350 0
## Correlation:
## (Intr)
## seedpred_presYes -0.956
##
## Standardized Within-Group Residuals:
## Min Q1 Med Q3 Max
## -22.4880970 -0.1451025 0.0000000 0.0000000 23.9565246
##
## Number of Observations: 6100
## Number of Groups:
## year sp %in% year

```

##

31

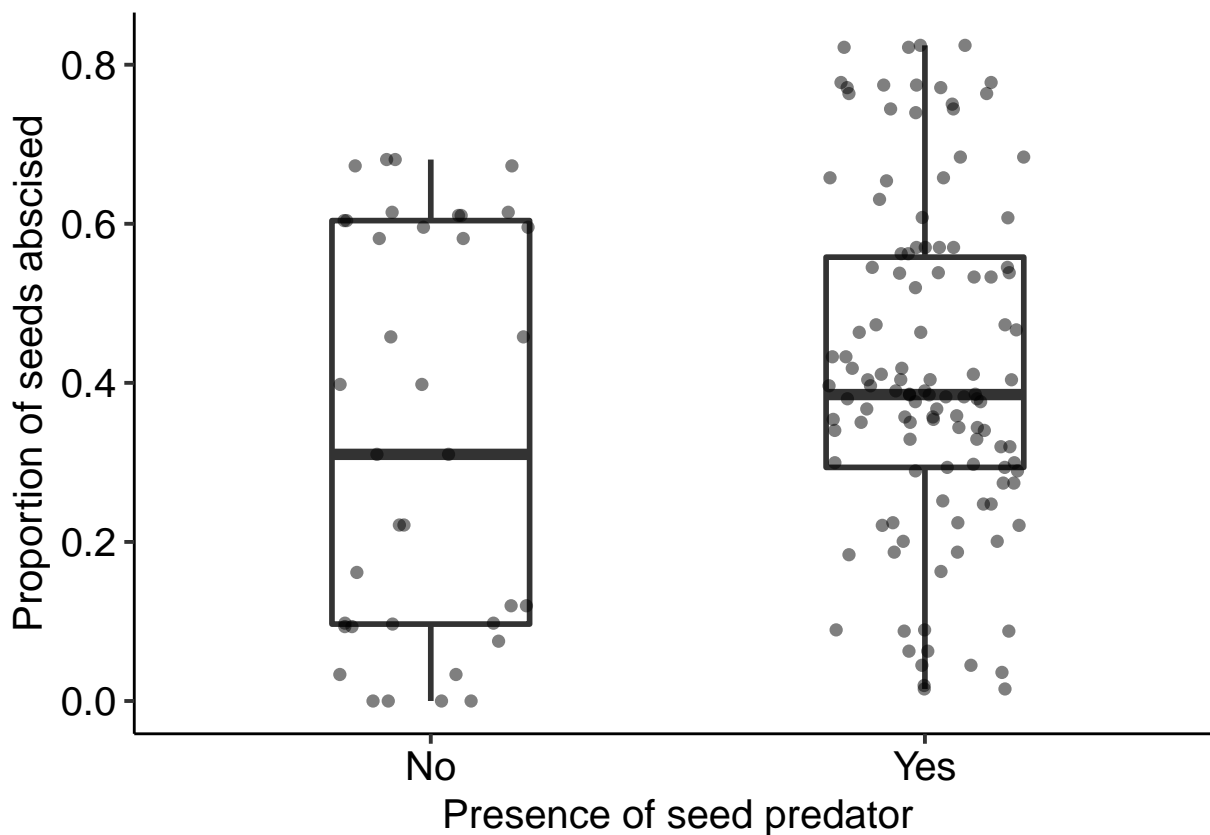
3782

## Check seedpred pres

*# a plot using the dataset with NAs removed*

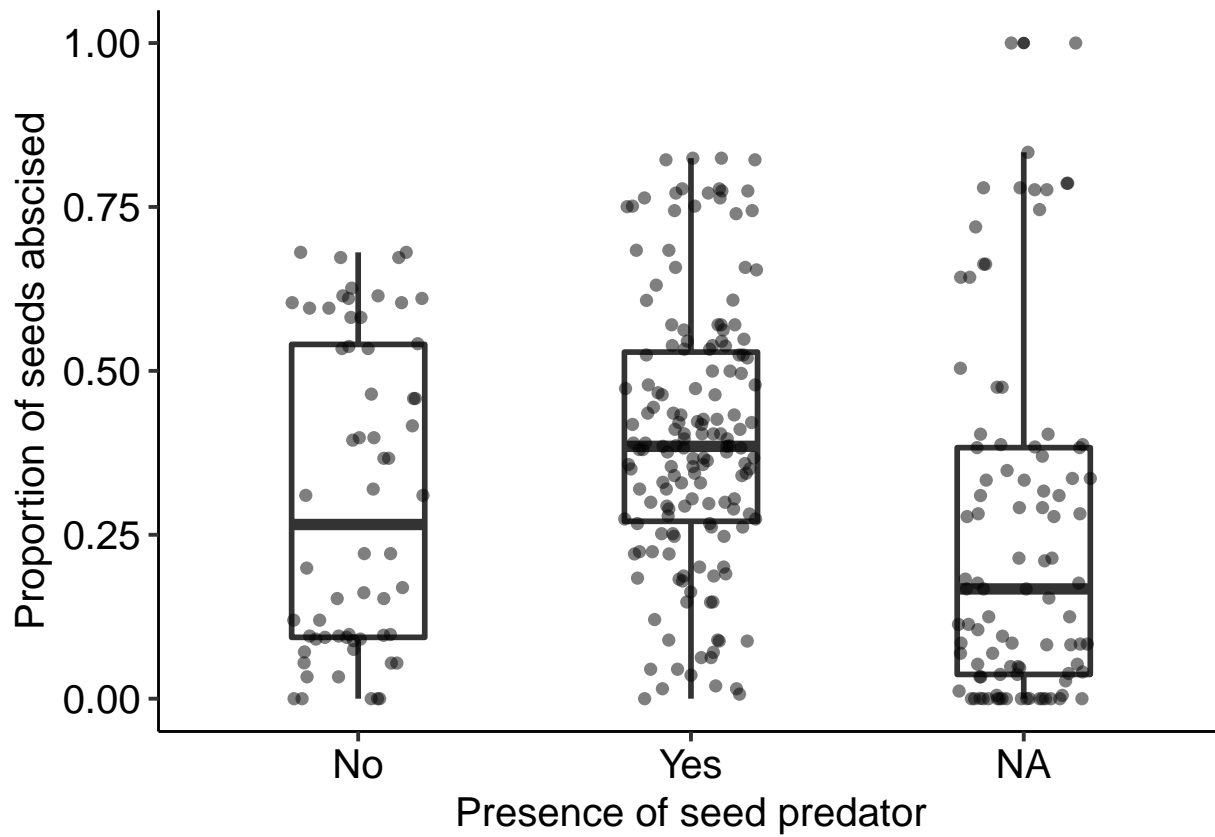
```
fruitDat.noNA <- fruitDat[complete.cases(fruitDat),
]
```

```
ggplot(fruitDat.noNA, aes(y = proportion_abcised,
  group = seedpred_pres, x = seedpred_pres)) + geom_boxplot(alpha = 0.8,
  position = position_dodge2(preserve = "single"),
  width = 0.4, size = 1) + geom_jitter(shape = 16,
  position = position_jitter(0.2), alpha = 0.5, size = 2) +
  ylab("Proportion of seeds abcised") + xlab("Presence of seed predator")
```



*# plot with fulldataset*

```
ggplot(fruitDat, aes(y = proportion_abcised, group = seedpred_pres,
  x = seedpred_pres)) + geom_boxplot(alpha = 0.8,
  position = position_dodge2(preserve = "single"),
  width = 0.4, size = 1) + geom_jitter(shape = 16,
  position = position_jitter(0.2), alpha = 0.5, size = 2) +
  ylab("Proportion of seeds abcised") + xlab("Presence of seed predator")
```



*# presence of seed predators corresponds to a  
# higher proportion of abscised seeds in both the  
# full dataset and the na's removed dataset*