ENV 710

Poisson regression

Duke | NICHOLAS SCHOOL OF THE ENVIRONMENT

roadmap

- Install/require packages: `performance`, `merTools`, `blmeco`, `glmmTMB`, `DHARMa`, `ggplot2`

concepts

- offset
- mixed effects model: fixed and random effects
- overdispersion
- zero-inflation

offsets – modeling rates

- model counts as a fraction of the total population or as a ratio of total effort
- model the rate at which events occur
- the adjustment term, *-log(t)* is called an offset

$$log(\mu/t) = \beta_0 + \beta_1 X_i$$

$$log(\mu) - log(t) = \beta_0 + \beta_1 X_i$$

$$\mu = t \cdot e^{(\beta_0 + \beta_1 X_i)}$$

```
glm(mu ~ factor(X), family = poisson, offset = log(t))
```

1 – owls

Abstract: Nestling begging behaviour may be an honest signal of need used by parents to adjust optimally both feeding rate and within-brood food allocation. Although several studies showed that mothers and fathers can be differentially responsive to nestling begging behaviour with one parent showing a stronger tendency to feed the offspring that beg the most, little information is yet available on whether offspring beg for food at different intensities from the mother than father. In the present study, we investigated in nestling barn owls whether the intensity of vocal begging behaviour in the presence of the mother and in the presence of the father is different. A difference is expected because reproductive tasks are divided between the sexes with fathers bringing more food items to the nest than mothers. The results show that although mothers transfer their prey item to one of the offspring more rapidly than fathers once in their nestbox, nestlings begged more intensely in the presence of their mother than in the presence of their father. To our knowledge, this is the first empirical evidence that offspring vocalize to different levels in the presence of their mother than in the presence of their father.

Roulin, A. and L. Bersier (2007) Nestling barn owls beg more intensely in the presence of their mother than in the presence of their father. Animal Behaviour 74 1099–1106.

I – owls

The data, taken from Zuur et al. (2009), quantify begging among owl nestlings in different nests prior to the arrival of a provisioning parent as a function of food treatment (deprived or satiated), the sex of the parent, and arrival time. The total number of calls from the nest is recorded.

```
summary(Owls) # from glmmTMB
Owls <- transform(Owls, Nest=reorder(Nest, NegPerChick),
                  NCalls=SiblingNegotiation,
                  FT=FoodTreatment)
```

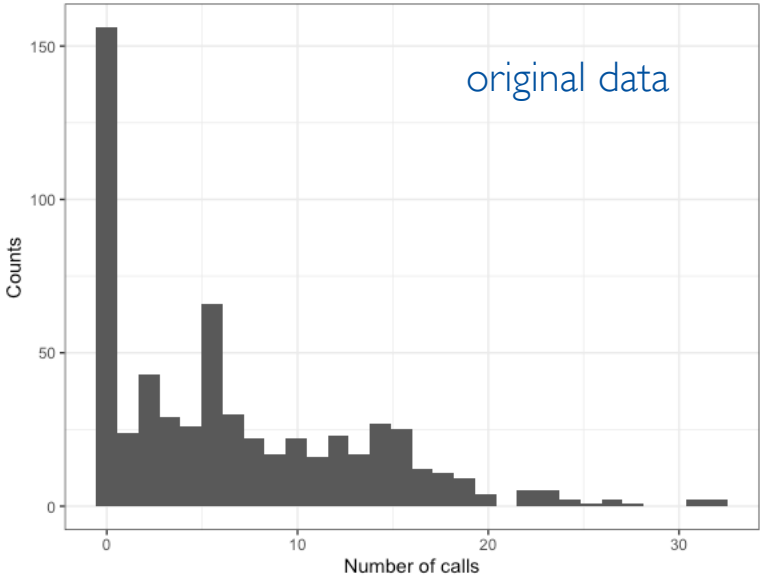
I – owls

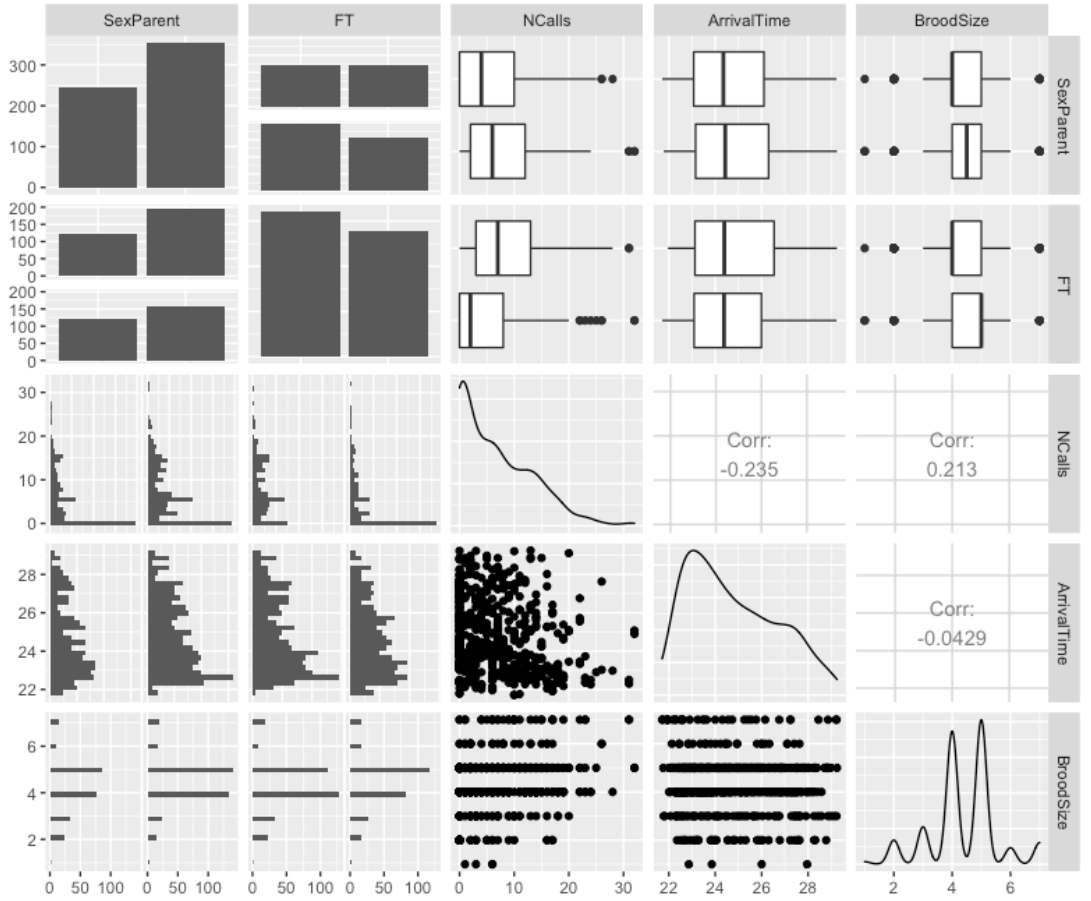
Does the number of calls per nestling depend on whether the nestlings are deprived or satiated with food or whether their mother or father are delivering the food? And does it depend on arrival time?

Model the mean number of calls per nestling.

- *NCalls*: number of calls from the nest
- *Nest*: factor describing individual nest locations
- *FT* (factor): food treatment, Deprived or Satiated
- *SexParent* (factor): sex of provisioning parent, Female or Male
- *ArrivalTime*: a numeric vector
- *BroodSize*: number of owlets in brood

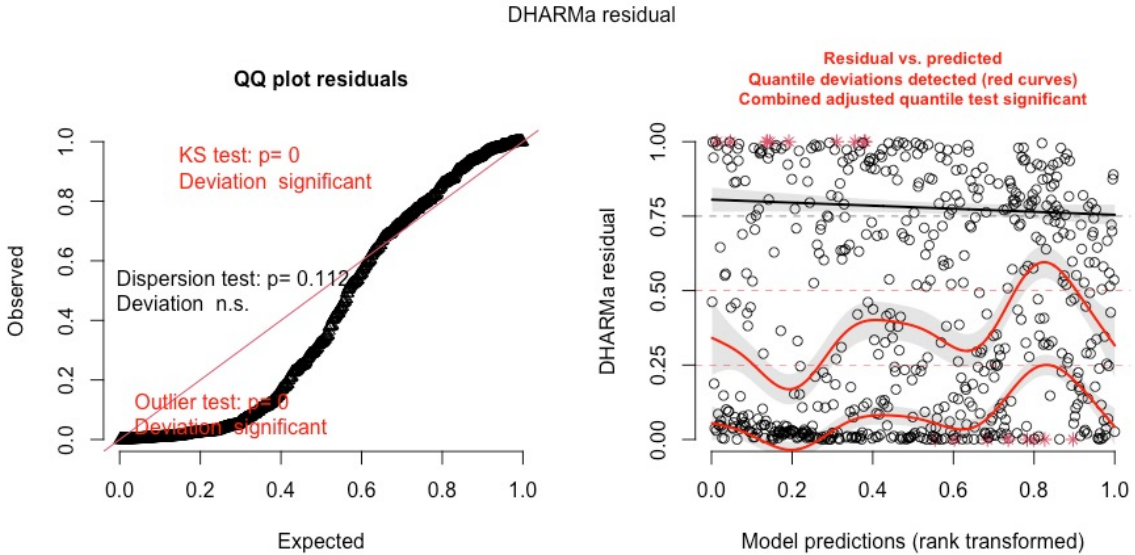
I – owls

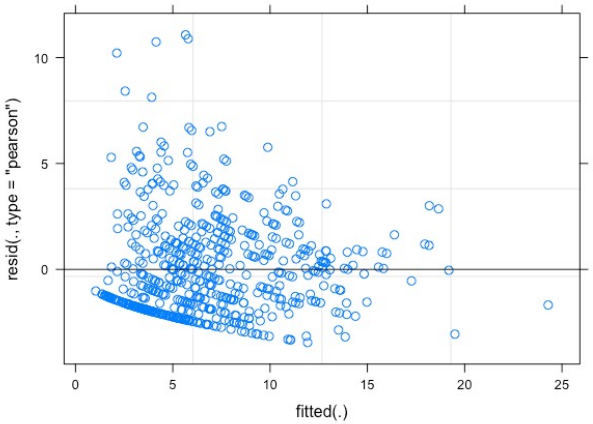

original data
Counts
Number of calls
0
50
100
150
0
10
20
30


SexParent
FT
NCalls
ArrivalTime
BroodSize
Corr:
-0.235
Corr:
0.213
Corr:
-0.0429

I – owls

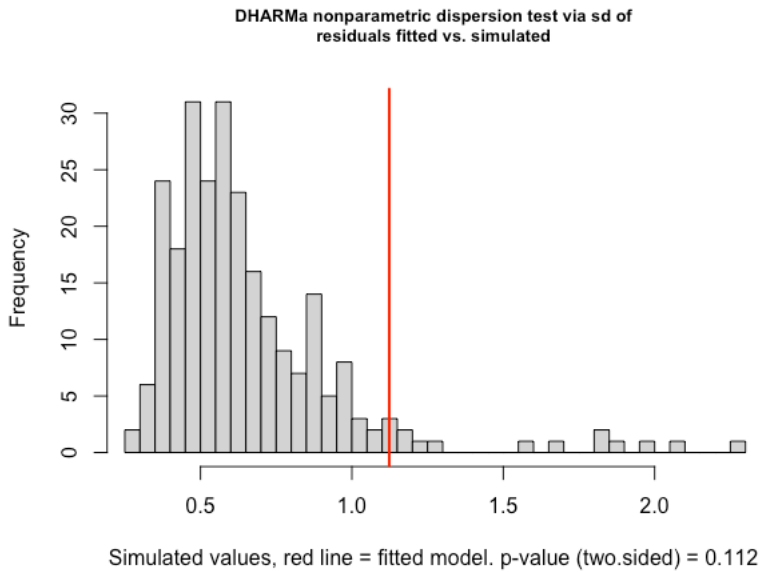
```
O0 <- glmer(NCalls ~ FT * SexParent + ArrivalTime + (1|Nest),
            offset = log(BroodSize),
            family = poisson, data = Owls)
plot(O0)
sim_res <- simulateResiduals(fittedModel = O0)
 plot(sim_res)
```





I – owls

```
O0 <- glmer(NCalls ~ FT * SexParent + ArrivalTime + (1|Nest),
            offset = log(BroodSize),
            family = poisson, data = Owls)
```




```
> testZeroInflation(O0)

DHARMa zero-inflation test via comparison to
expected zeros with simulation under H0 =
fitted model

data:  simulationOutput
ratioObsSim = 9.7573, p-value < 2.2e-16
alternative hypothesis: two.sided

> testDispersion(O0)

DHARMa nonparametric dispersion test via sd of
residuals fitted vs. simulated

data:  simulationOutput
dispersion = 1.7136, p-value = 0.112
alternative hypothesis: two.sided
```

I – owls

Zero-inflation is the situation when more zeros appear in the observation than expected under the fitted model.

There are two zero-generating processes.:

1. zeros generated by the first (binary) distribution are "structural";
2. zeros generated by the second (count) distribution are "random".

Can use (at least) two approaches:

1. Hurdle model
2. Zero-inflated Poisson model (ZIP)

model number of calls conditional on owlets calling

```
Owls1 <- Owls[Owls$NCalls > 0,]
```

I – owls

```
o1 <- glmer(NCalls ~ FT * SexParent + (1|Nest),
              offset = log(BroodSize),
              family = poisson, data = Owls1)
summary(o1)
```

```
Random effects:
 Groups Name        Variance Std.Dev.
 Nest   (Intercept) 0.1211   0.348
Number of obs: 443, groups:  Nest, 27

Fixed effects:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                2.824271   0.236372  11.948  < 2e-16 ***
FTSatiated                -0.285860   0.058370  -4.897 9.71e-07 ***
SexParentMale             -0.070956   0.045784  -1.550    0.121
ArrivalTime               -0.079685   0.009167  -8.693  < 2e-16 ***
FTSatiated:SexParentMale   0.105018   0.071763   1.463    0.143
```

```
o2 <- update(o1, .~. -FT:SexParent)
summary(o2)
```

```
Random effects:
 Groups Name        Variance Std.Dev.
 Nest   (Intercept) 0.1208   0.3476
Number of obs: 443, groups:  Nest, 27

Fixed effects:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    2.791684   0.235299  11.864  < 2e-16 ***
FTSatiated    -0.219975   0.037013  -5.943  2.8e-09 ***
SexParentMale -0.029541   0.036224  -0.816    0.415
ArrivalTime   -0.079426   0.009163  -8.668  < 2e-16 ***
```

I – owls

```
o3 <- update(o2, .~. -SexParent)
summary(o3)
```

```
Random effects:
 Groups Name        Variance Std.Dev.
 Nest   (Intercept) 0.1206   0.3473
Number of obs: 443, groups:  Nest, 27

Fixed effects:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.78042    0.23502  11.831  < 2e-16 ***
FTSatiated  -0.21772    0.03692  -5.897  3.7e-09 ***
ArrivalTime -0.07975    0.00916  -8.706  < 2e-16 ***
```

```
require(performance)
check_overdispersion(o3)

# Overdispersion test

 dispersion ratio =    3.597
 Pearson's Chi-Squared = 1579.003
 p-value =  < 0.001
```

1 – owls

```
o4 <- update(o3, .~. + (1|N))
summary(o4)
```

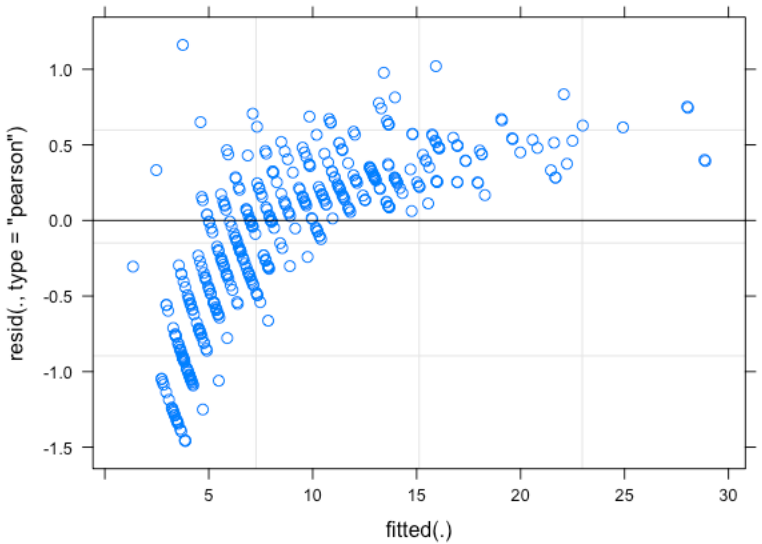
```
Random effects:
 Groups Name        Variance Std.Dev.
 N      (Intercept) 0.33432  0.5782
 Nest   (Intercept) 0.06756  0.2599
Number of obs: 443, groups:  N, 443; Nest, 27

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.51905    0.44374   5.677 1.37e-08 ***
FTSatiated  -0.25489    0.07346  -3.470 0.000521 ***
ArrivalTime -0.07507    0.01783  -4.210 2.56e-05 ***
```

corrected overdispersion by adding an observation-level random effect, but residuals don't look good

```
check_overdispersion(o4)
```

```
# Overdispersion test

 dispersion ratio =   0.304
  Pearson's Chi-Squared = 133.269
              p-value =       1
```

```
plot(o4)
```

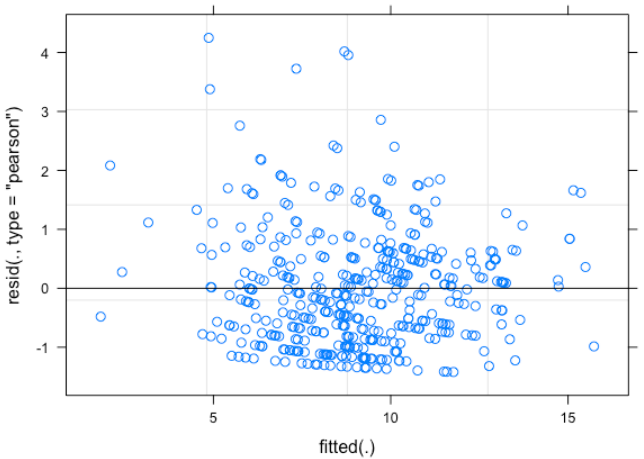


I – owls

corrected overdispersion by using a negative binomial model

```
o5 <- glmer.nb(NCalls ~ FT + (1|Nest) +
        offset(log(BroodSize)), data = Owls1)
 summary(o5)

Family: Negative Binomial(3.0291)  ( log )
Formula: NCalls ~ factor(FT) + ArrivalTime + (1 | Nest) +
         offset(log(BroodSize))
   Data: Owls1

Random effects:
 Groups Name        Variance Std.Dev.
 Nest   (Intercept) 0.06377  0.2525
Number of obs: 443, groups:  Nest, 27

Fixed effects:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          2.51856    0.43243   5.824 5.74e-09 ***
factor(FT)Satiated -0.19698    0.07247  -2.718  0.00657 **
ArrivalTime        -0.06958    0.01742  -3.993 6.51e-05 ***
```

```
plot(o5)
```




```
> with(Owls1, bartlett.test(NCalls/BroodSize ~ FT))

        Bartlett test of homogeneity of variances

data:  NCalls/BroodSize by FT
Bartlett's K-squared = 0.054589, df = 1, p-value = 0.8153
```

I – owls

```
> AIC(o0,o1,o2, o3, o4, o5)

   df      AIC
o0  6 3329.110
o1  6 3329.110
o2  5 3329.246
o3  4 3327.908
o4  5 2769.352
o5  5 2759.632
```

```
> lrtest(o1, o2, o3, o4, o5)
Likelihood ratio test

Model 1: NCalls ~ FT * SexParent + ArrivalTime + (1 | Nest)
Model 2: NCalls ~ FT + SexParent + ArrivalTime + (1 | Nest)
Model 3: NCalls ~ FT + ArrivalTime + (1 | Nest)
Model 4: NCalls ~ FT + ArrivalTime + (1 | Nest) + (1 | N)
Model 5: NCalls ~ factor(FT) + ArrivalTime + (1 | Nest) +
         offset(log(BroodSize))
  #Df  LogLik Df    Chisq Pr(>Chisq)
1   6 -1658.6
2   5 -1659.6 -1   2.1360     0.1439
3   4 -1660.0 -1   0.6618     0.4159
4   5 -1379.7  1 560.5562     <2e-16 ***
5   5 -1374.8  0   9.7199     <2e-16 ***
```

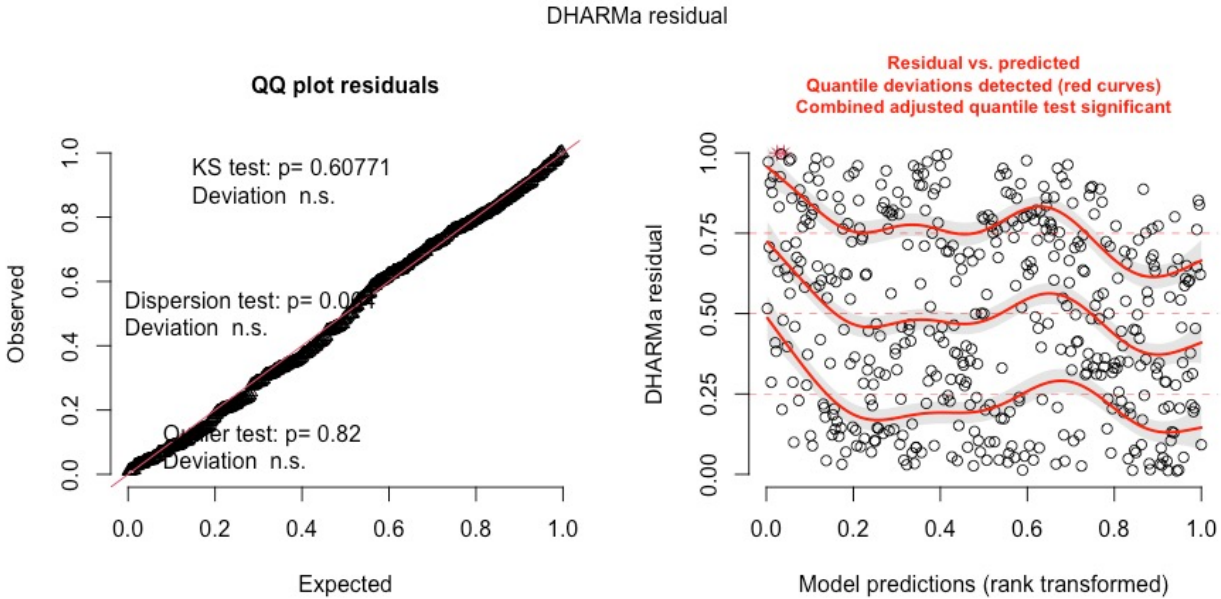
o5 – negative binomial model seems to be the best model…

I – owls

```
> res5 <- simulateResiduals(fittedModel = o5)
> plot(res5)
```

```
> testZeroInflation(res5)
DHARMa zero-inflation test via comparison to expected
zeros with simulation
                under H0 = fitted model

data:  simulationOutput
ratioObsSim = 0, p-value < 2.2e-16
alternative hypothesis: two.sided
```

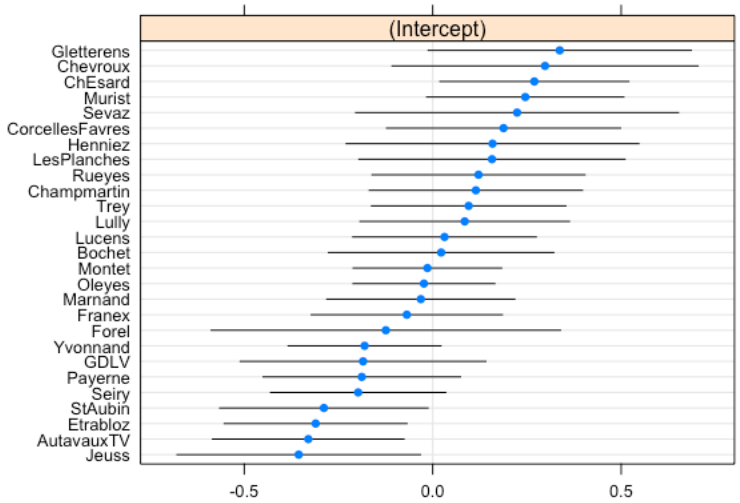
```
> testDispersion(o5)

DHARMa nonparametric dispersion test via sd of residuals
fitted vs. simulated

data:  simulationOutput
dispersion = 0.71177, p-value = 0.064
alternative hypothesis: two.sided
```

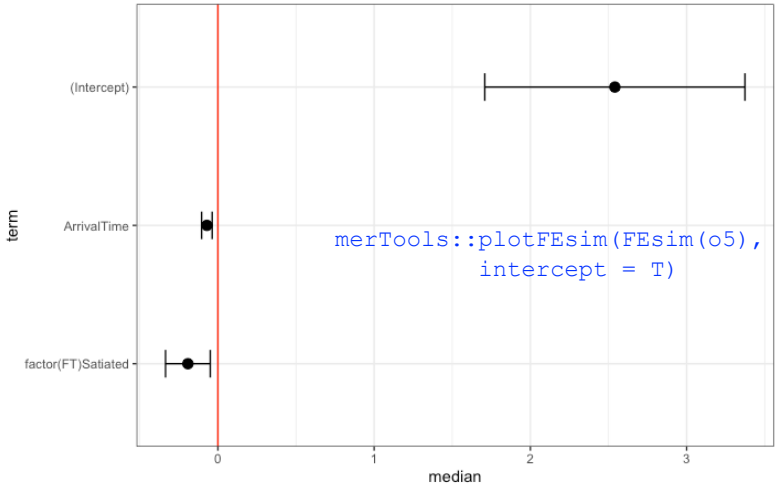


I – owls

```
> MuMIn::r.squaredGLMM(o5)
                R2m       R2c
delta     0.04741312 0.1671464
lognormal 0.05536006 0.1951619
trigamma  0.03914937 0.1380141
```

```
dotplot(ranef(o5))
```






```
merTools::plotFEsim(FEsim(o5),
                    intercept = T)
```

I – owls

```
summary(o5)
```

```
Random effects:
 Groups Name        Variance Std.Dev.
 Nest   (Intercept) 0.06377  0.2525
Number of obs: 443, groups:  Nest, 27

Fixed effects:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)         2.51856    0.43243   5.824 5.74e-09 ***
factor(FT)Satiated -0.19698    0.07247  -2.718  0.00657 **
ArrivalTime        -0.06958    0.01742  -3.993 6.51e-05 ***
------
```

estimate for 'deprived'

```
exp(fixef(o5)[1])
```

```
(Intercept)
   12.41071
```

baseline rate of owl calls for chicks deprived of food at 0 arrival time!

estimate for 'satiated'

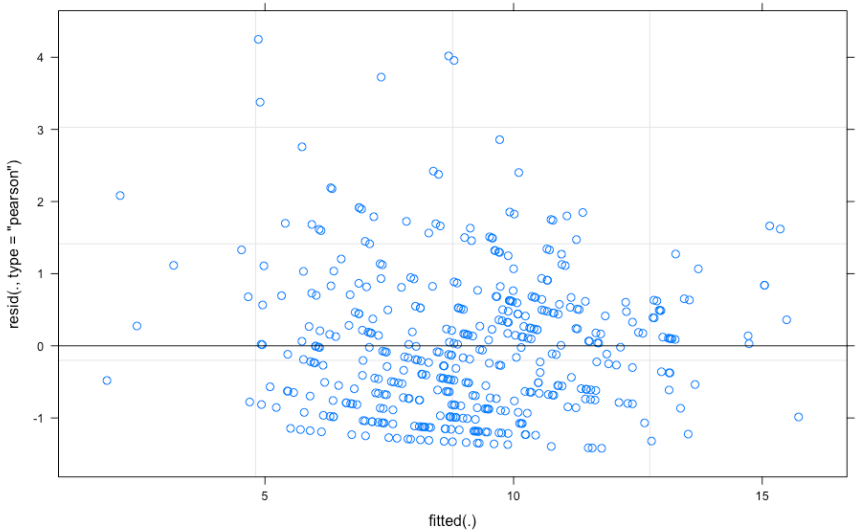
```
> exp(fixef(o5)[2])
```

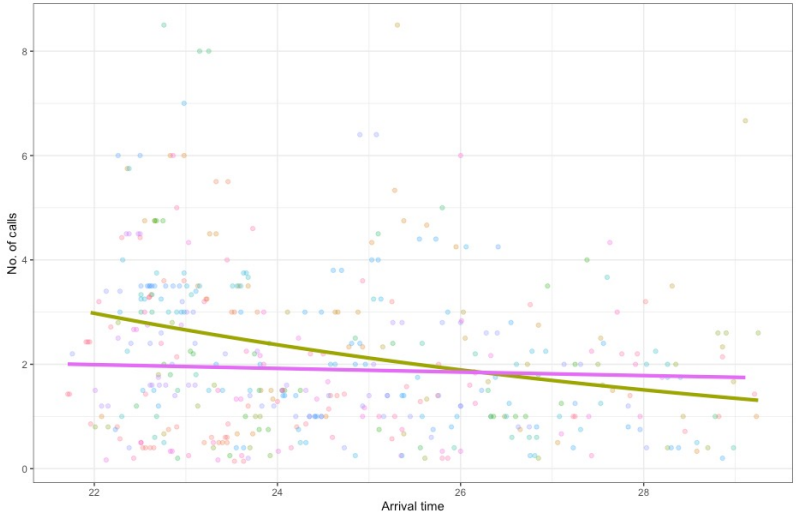
```
factor(FT)Satiated
         0.8212089
```

calls of satiated chicks are 0.82 times that of deprived owls, or ~ 18% less

average number of calls for 'satiated' at mean arrival time of 24.6 s

```
> exp(fixef(o5)[1] + fixef(o5)[2]) +
      fixef(o5)[3]*mean(Owls1$ArrivalTime)
(Intercept)
   8.479667
```






```
ggplot(Owls1, aes(x = ArrivalTime, y = NCalls/BroodSize, group = factor(FT),
                  col = factor(FT))) +
  geom_point(aes(col= factor(Nest)), alpha = 0.3) +
    geom_smooth(method='glm', method.args = list(family = 'poisson'), se = F,
                size=1.5) +
    ylab("No. of calls") + xlab("Arrival time") +
    theme_bw() + theme(legend.position = "none")
```

Questions?