Ch1\_Blackbear\_modelDistribution

Erin Tattersall

May 8, 2018

Modelling detections rates between November 2015 and April 2018 for black bear.  
Deciding on most appropriate distribution for response variable.

Proportion of zeros in data

sum(det$Blackbear==0, na.rm = TRUE)/nrow(det)

## [1] 0.825

82.5% of the data is zeroes –> likely zero-inflated.

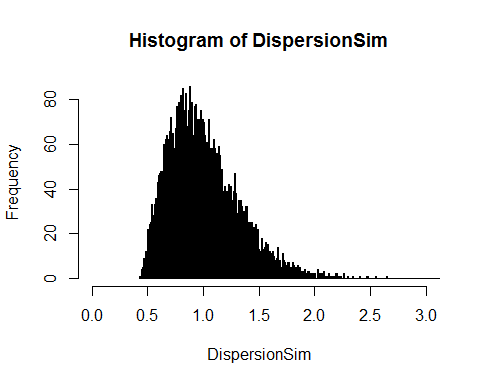
### Fitting a basic poisson GLM and checking overdispersion, using global model (doesn’t yet include SnowDays)

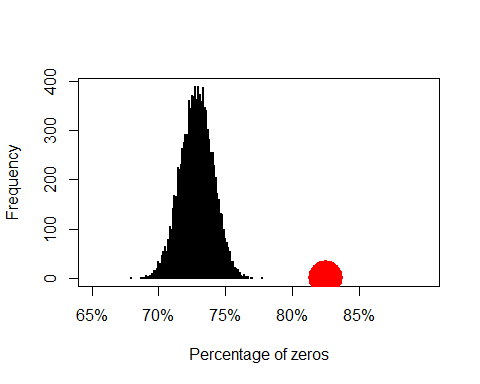
glm1 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays, data = det, family = poisson)  
# Residuals and overdispersion  
E1 <- resid(glm1, type="pearson")  
N <- nrow(det)  
p <- length(coef(glm1))  
sum(E1^2)/(N-p)

## [1] 1.882168

Dispersion value of 1.88 indicates data is probably overdispersed

Simulating data to check probability of calculated dispersion.

 Histogram indicates that an overdispersion statistic of 1.88 is toward the high edge of the likely distribution of dispersion statistics for Poisson distributed response variables, suggesting that Blackbear data is probably overdispersed.

Comparing the proportion of zeros in data to simulated zeros from model shows that Blackbear data has more zeroes than would be expected in a Poisson GLM  This presents a case for using zero-inflated models, which can be verified with model selection of GLMMs ## Model selection: choosing model form and distribution

Comparing the same GLMM (including random effects of Site and Month) modeled as a poisson, nb, ZIP and ZINB (with nbinom1 and nbinom2 differing in how variance changes with mean) yields:

## Model comparisons of distributions and zero-inflation  
glm1 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays + (1| Site) + (1|Month), data = det, family = poisson)  
glm2 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays + (1| Site) + (1|Month), data = det, family = nbinom1(link= "log")) #warning matrix not positive definite  
glm3 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays + (1| Site) + (1|Month), data = det, family = nbinom2(link = "log"))  
glm4 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays + (1| Site) + (1|Month), zi = ~1, data = det, family = poisson)  
glm5 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays +(1| Site) + (1|Month), zi = ~1, data = det, family = nbinom1(link= "log")) ## warning 'matrix not positive definite' and 'lgamma value out of range'  
glm6 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays + (1| Site) + (1|Month), zi = ~1, data = det, family = nbinom2(link= "log")) ## warning 'matrix not positive definite'

## dLogLik dAIC df weight  
## Nbinom2 21.1 0.0 12 0.552   
## ZINB2 21.1 2.0 13 0.203   
## Nbinom1 19.9 2.3 12 0.174   
## ZINB1 20.0 4.1 13 0.071   
## ZIP 9.3 23.6 12 <0.001  
## Poisson 0.0 40.1 11 <0.001

Nbinom1, ZINB1, and ZINB2 displayed a warning when modeled, but parameter estimates + s.e. obtained, so included in model selection.  
Nbinom2 model’s summary output is:

## Family: nbinom2 ( log )  
## Formula:   
## Blackbear ~ Treatment + LineWidth + LD750 + VegHt + low500 +   
## ActiveDays + (1 | Site) + (1 | Month)  
## Data: det  
##   
## AIC BIC logLik deviance df.resid   
## 939.3 995.9 -457.7 915.3 814   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## Site (Intercept) 0.8997 0.9485   
## Month (Intercept) 0.5912 0.7689   
## Number of obs: 826, groups: Site, 59; Month, 7  
##   
## Overdispersion parameter for nbinom2 family (): 1.46   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.27359 1.33733 -4.691 2.72e-06 \*\*\*  
## TreatmentHumanUse 0.77037 0.53540 1.439 0.15019   
## TreatmentNatRegen -0.19146 0.52578 -0.364 0.71576   
## TreatmentSPP 0.30022 0.43541 0.690 0.49050   
## LineWidth -0.02351 0.12569 -0.187 0.85165   
## LD750 -0.05218 0.21089 -0.247 0.80456   
## VegHt 0.48313 0.18505 2.611 0.00903 \*\*   
## low500 -0.79724 0.86702 -0.920 0.35783   
## ActiveDays 0.17240 0.02378 7.251 4.13e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ZINB model output:

## Family: nbinom2 ( log )  
## Formula:   
## Blackbear ~ Treatment + LineWidth + LD750 + VegHt + low500 +   
## ActiveDays + (1 | Site) + (1 | Month)  
## Zero inflation: ~1  
## Data: det  
##   
## AIC BIC logLik deviance df.resid   
## 941.3 1002.6 -457.7 915.3 813   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## Site (Intercept) 0.8997 0.9485   
## Month (Intercept) 0.5912 0.7689   
## Number of obs: 826, groups: Site, 59; Month, 7  
##   
## Overdispersion parameter for nbinom2 family (): 1.46   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.27359 1.33733 -4.691 2.72e-06 \*\*\*  
## TreatmentHumanUse 0.77037 0.53540 1.439 0.15019   
## TreatmentNatRegen -0.19145 0.52578 -0.364 0.71576   
## TreatmentSPP 0.30022 0.43541 0.690 0.49050   
## LineWidth -0.02351 0.12569 -0.187 0.85164   
## LD750 -0.05218 0.21089 -0.247 0.80457   
## VegHt 0.48313 0.18505 2.611 0.00903 \*\*   
## low500 -0.79724 0.86702 -0.920 0.35783   
## ActiveDays 0.17240 0.02378 7.251 4.13e-13 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Zero-inflation model:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -17.75 4293.38 -0.004 0.997

Top model Nbinom2 did not result in any warnings, but both ZINB models did. Proceed with ZINB with caution.  
Active Days affects the probability of observing a zero in data –> should be included in ZI model. It could be argued that it should NOT be included in conditional. I will test both

glm7 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + ActiveDays +(1| Site) + (1|Month), zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))  
glm8 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + (1| Site) + (1|Month), zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))  
summary(glm8)

## Family: nbinom2 ( log )  
## Formula:   
## Blackbear ~ Treatment + LineWidth + LD750 + VegHt + low500 +   
## (1 | Site) + (1 | Month)  
## Zero inflation: ~ActiveDays  
## Data: det  
##   
## AIC BIC logLik deviance df.resid   
## 936.3 997.7 -455.2 910.3 813   
##   
## Random effects:  
##   
## Conditional model:  
## Groups Name Variance Std.Dev.  
## Site (Intercept) 0.8524 0.9233   
## Month (Intercept) 0.5439 0.7375   
## Number of obs: 826, groups: Site, 59; Month, 7  
##   
## Overdispersion parameter for nbinom2 family (): 1.41   
##   
## Conditional model:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -0.92605 1.09151 -0.848 0.3962   
## TreatmentHumanUse 0.71979 0.52706 1.366 0.1720   
## TreatmentNatRegen -0.30710 0.51555 -0.596 0.5514   
## TreatmentSPP 0.27907 0.42666 0.654 0.5131   
## LineWidth -0.01704 0.12342 -0.138 0.8902   
## LD750 -0.05131 0.20749 -0.247 0.8047   
## VegHt 0.41483 0.18796 2.207 0.0273 \*  
## low500 -0.85278 0.85507 -0.997 0.3186   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Zero-inflation model:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 332.19 150040.59 0.002 0.998  
## ActiveDays -31.89 14977.19 -0.002 0.998

## dLogLik dAIC df weight  
## ZINB2-AD1 23.5 0.0 13 0.710   
## Nbinom2 21.1 3.0 12 0.160   
## ZINB2 21.1 5.0 13 0.059   
## Nbinom1 19.9 5.3 12 0.050   
## ZINB1 20.0 7.1 13 0.021   
## ZIP 9.3 26.6 12 <0.001  
## Poisson 0.0 43.1 11 <0.001  
## ZINB2-AD2 NA NA 14 NA

Adding ActiveDays to ZI model resulted in a model without convergence issues that had top model weight I will therefore model Blackbear data as zero-inflated nbinom2 with ActiveDays in the ZI model.

# Model hypotheses  
## Finding random structure

r0 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500, zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))

## Warning in fitTMB(TMBStruc): Model convergence problem; singular  
## convergence (7). See vignette('troubleshooting')

rSite <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + (1| Site), zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))   
rMonth <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + (1|Month), zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))

## Warning in fitTMB(TMBStruc): Model convergence problem; singular  
## convergence (7). See vignette('troubleshooting')

r2 <- glmmTMB(Blackbear~ Treatment + LineWidth + LD750 + VegHt + low500 + (1| Site) + (1|Month), zi = ~ActiveDays, data = det, family = nbinom2(link= "log"))

## Warning in fitTMB(TMBStruc): Model convergence problem; singular  
## convergence (7). See vignette('troubleshooting')

ICtab(r0, rSite, rMonth, r2, type= "AIC", weights = TRUE, delta = TRUE, logLik = TRUE, sort=TRUE)

## dLogLik dAIC df weight  
## r2 40.7 0.0 13 1   
## rSite 19.3 40.9 12 <0.001  
## rMonth 17.4 44.7 12 <0.001  
## r0 0.0 77.5 11 <0.001

Models testing random structure resulted in convergence issues, despite having worked with similar structure previously. I will continue modelling with 2 random effects for consistency and because it makes sense (justification stated elsewhere)