Sawyer Beetle Trapping Study

Comparing lure effectiveness and models

Project aims

- To determine if added chemical or height of trap (or both) have a significant impact on beetle capture

- To determine if modeling using GLM count models outperforms or solves any problems with log transformed + ANOVA approach

Setup

- Variables: basic chemical lures (m-) or basic + monochamol (m+) at canopy or ground height
- 4 traps per block, one of each unique combination
- 6 blocks in transect
- Ran transect at 4 time periods
- Counted Monochamus beetle individuals in each trap at end of time periods

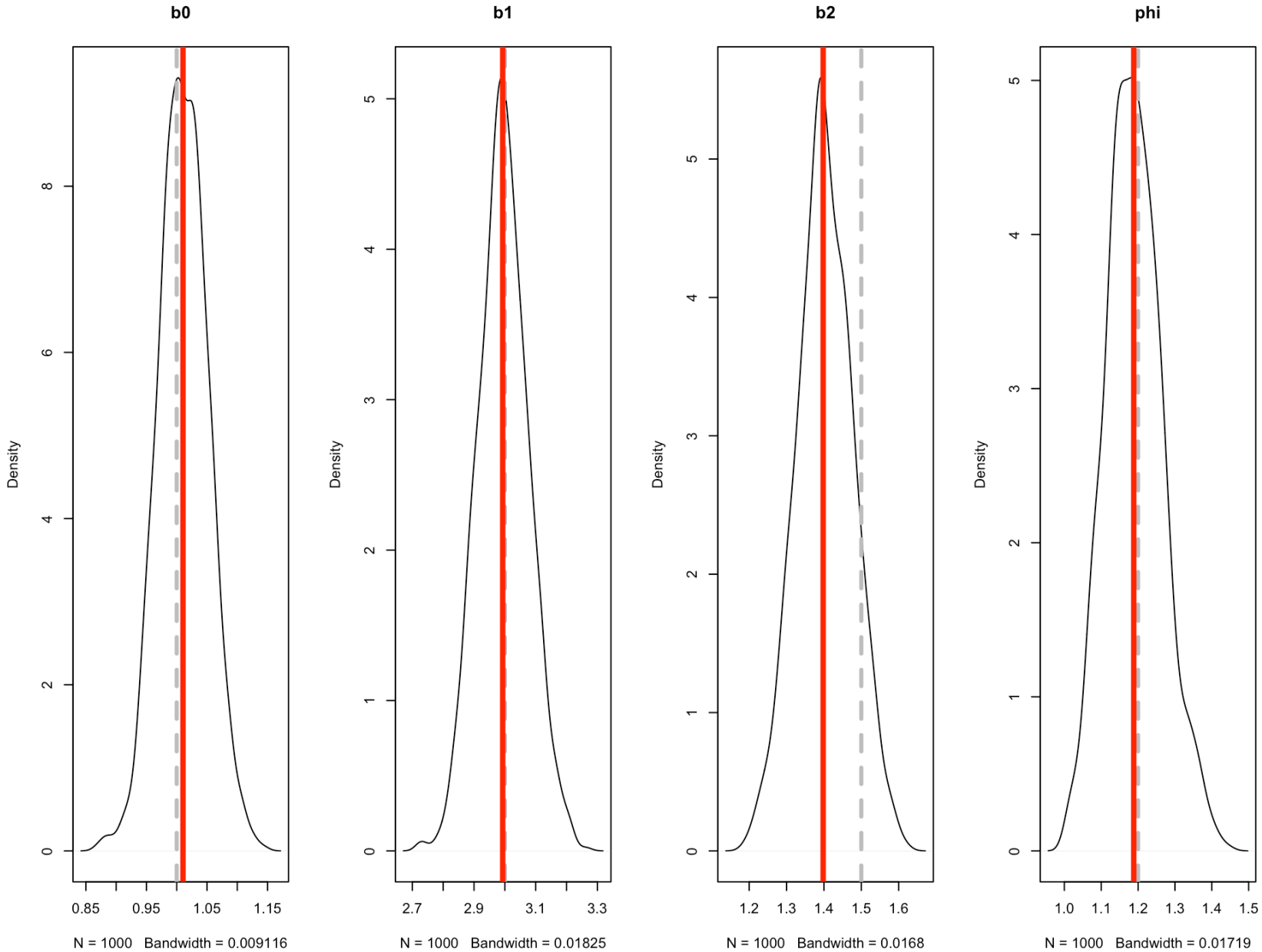
Data

- Counted Monochamus spp. beetle individuals in each trap at end of time periods
- 24 traps counted per time, 96 total (24 of each unique condition)
- Total 536 beetles
- Mean beetles/trap for each condition:
  - Hi+ 4.25
  - Hi- 2.08
  - Lo+ 12
  - Lo- 4

Fake data

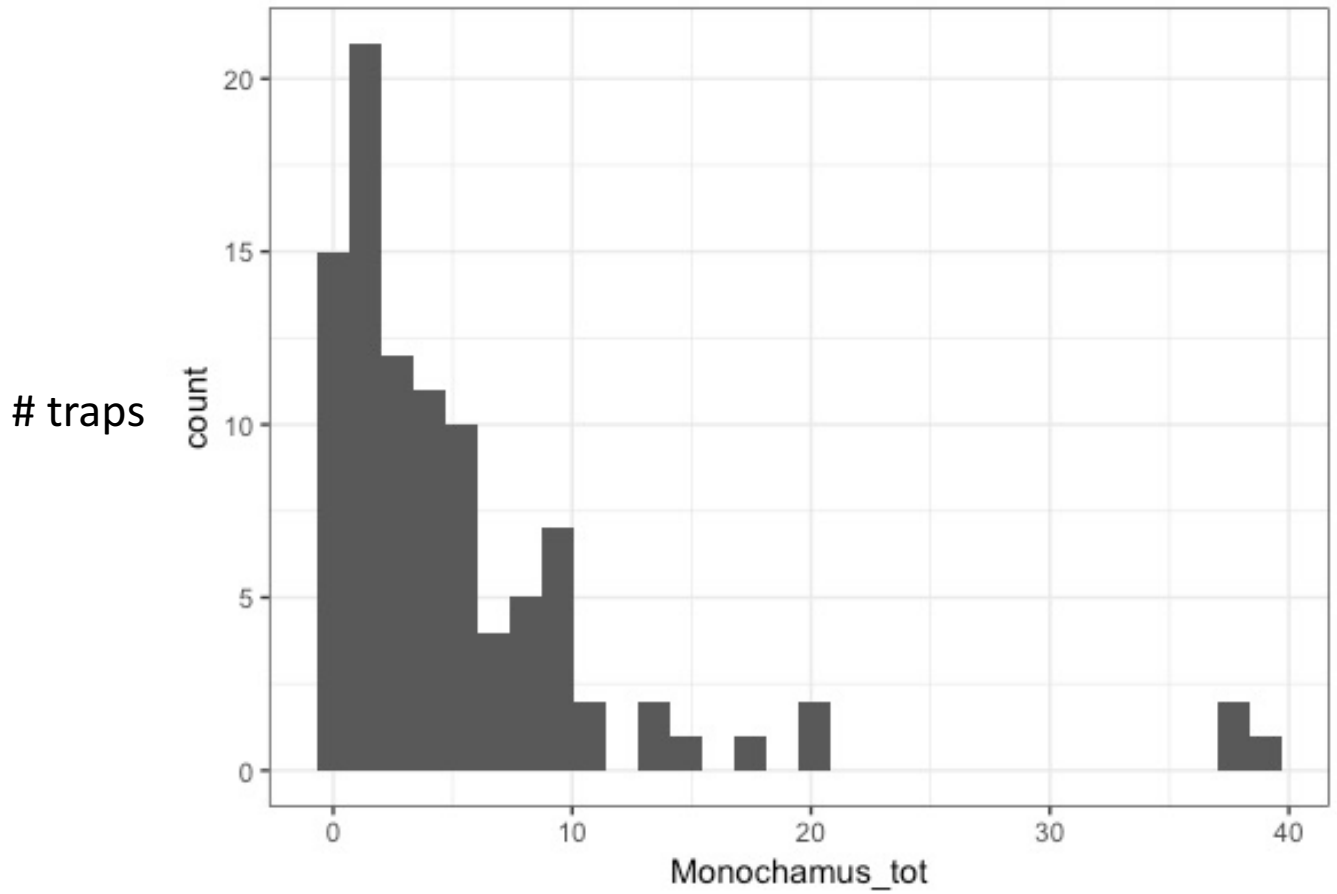
- ## simulate fake data using the negative binomial, phi=theta
- set.seed(123)
- N <- 1000
- phi <- 1.2
- b0 <- 1
- b1 <- 3
- b2 <-1.5
- b3 <- 1
- x1 <- rbinom(N, 1, 0.5)
- x2 <- rbinom(N, 1, 0.5)
- y_2var <- rnbinom(N, size = phi, mu = exp(b0 + (x1 - mean(x1)) * b1 + (x2 - mean(x2)) * b2))
- plot(x, y_2var)

Comparing expectations to results for fake data
Neg binomial seems to work ok with heavy sampling



Gray-true
Red-derived
from simulations

Modeling many low values with moderate variance around the mean and some extremes



Model coefficient outputs

Model	Intercept	SD	height low	SD	lures plain	SD
Neg Binomial 4 var	0.1	0.2	0.9	0.1	-0.9	0.1
Poisson 4 var	0.1	0.2	0.9	0.1	-1	0.1

Leave one out model comparison

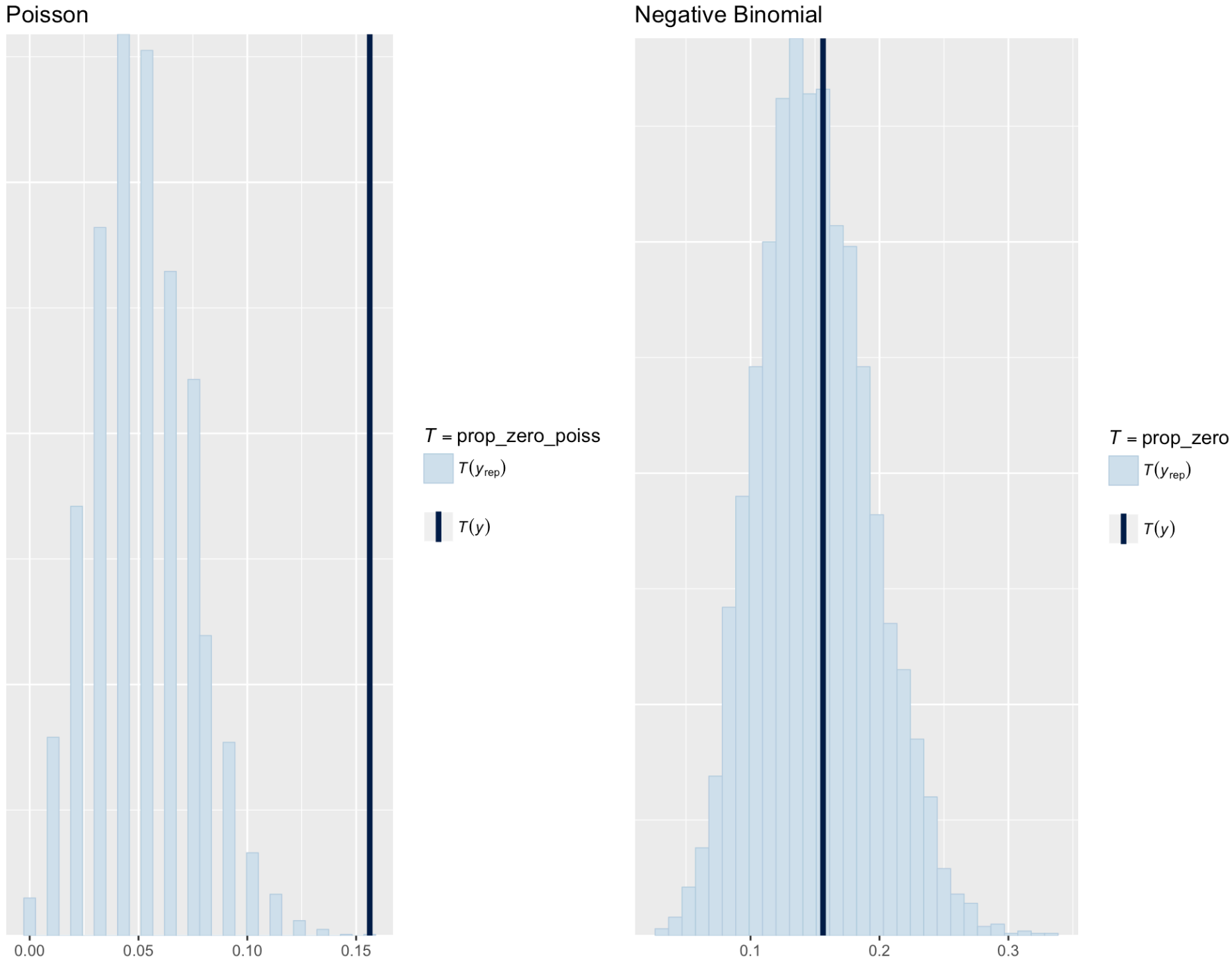
- loo R package (Gelman) for bayesian l.o.o. comparisons

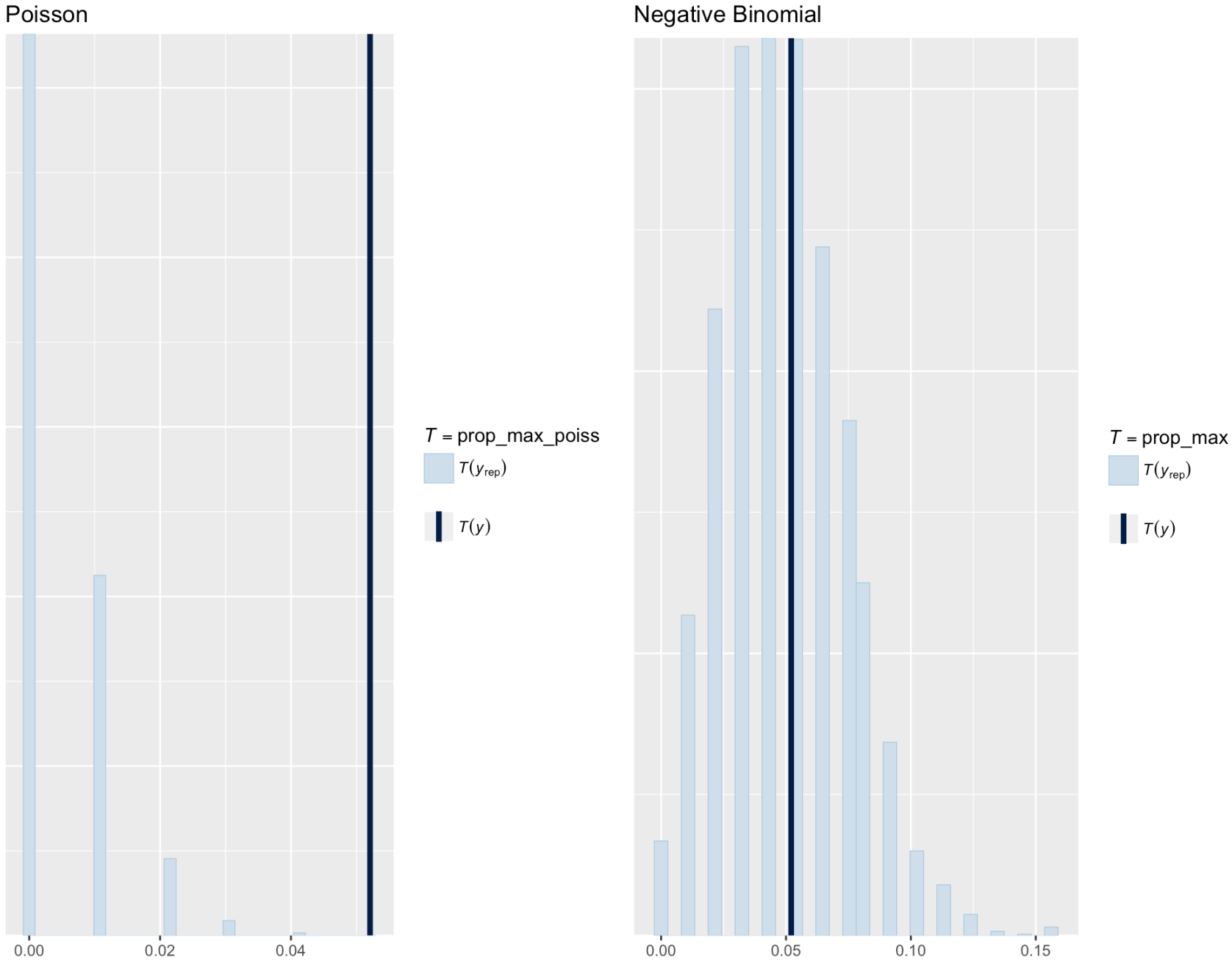
- Compared poisson and negbinomial models with height and lures
  - added date and block predictors for four variable model
- Best height + lures model is negative binomial
  - predictive density_diff = 106.4 se 30.8
- Using all 4 variables there is no model preference
  - predictive density_diff = .2 se 1.5

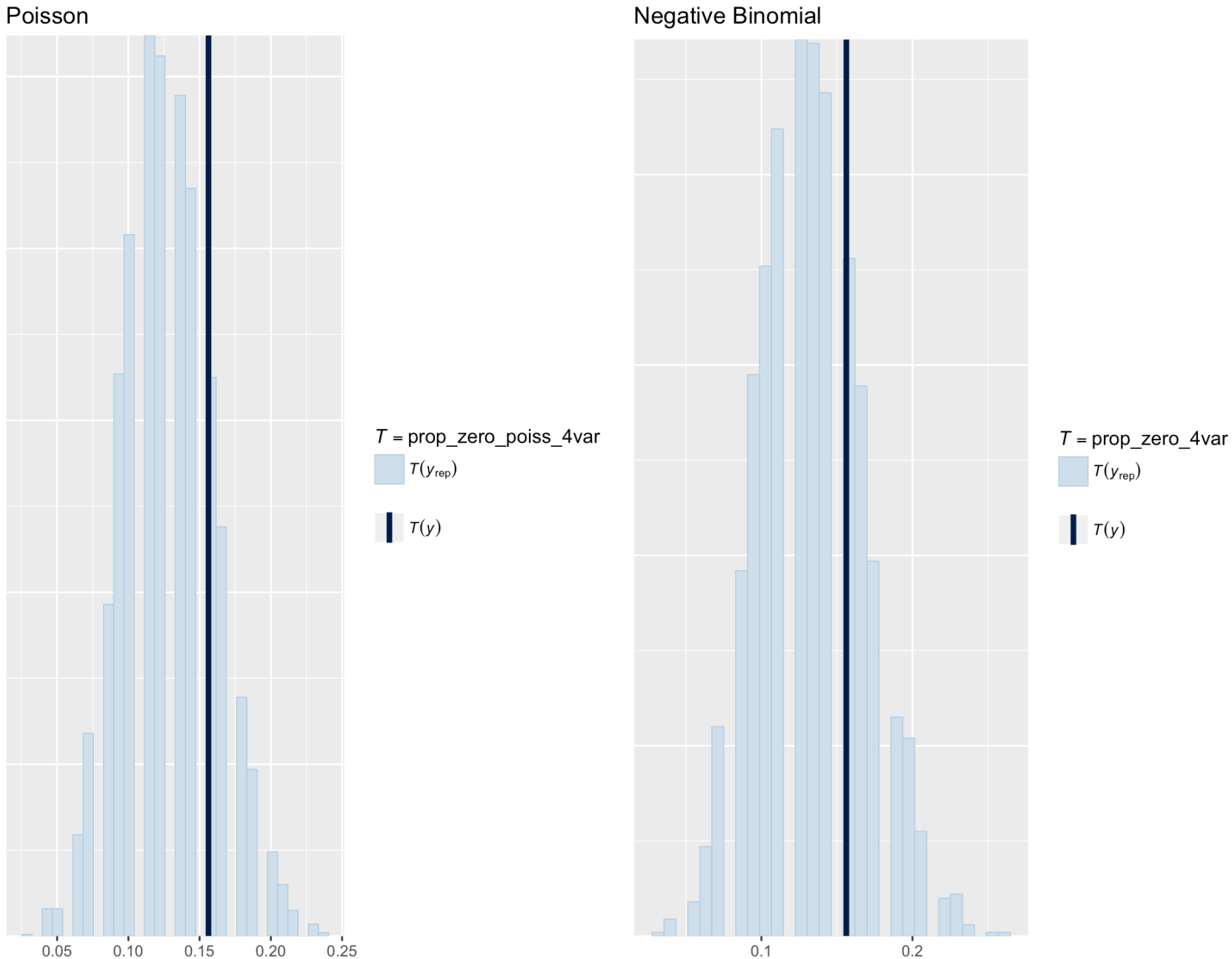
- There is potentially compensation for zeroes in low catch dates when more parameters are included?

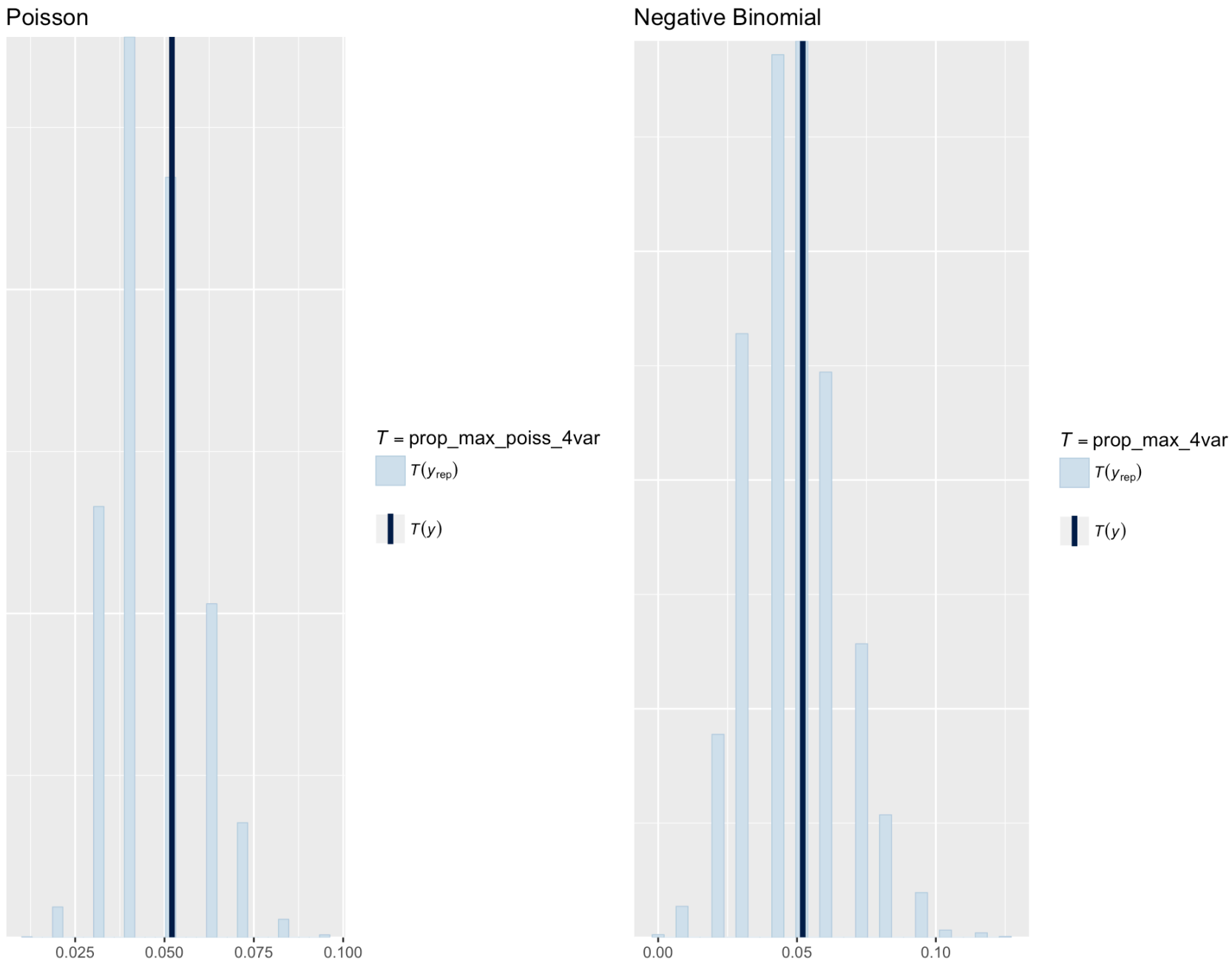
Posterior Predictive checks

Two variable and four variable Bayesian GLM models


Poisson
0.00
0.05
0.10
0.15
T = prop_zero_poiss
T(y_rep)
T(y)
Negative Binomial
0.1
0.2
0.3
T = prop_zero
T(y_rep)
T(y)


Poisson
0.00
0.02
0.04
T = prop_max_poiss
T(y_rep)
T(y)
Negative Binomial
0.00
0.05
0.10
0.15
T = prop_max
T(y_rep)
T(y)

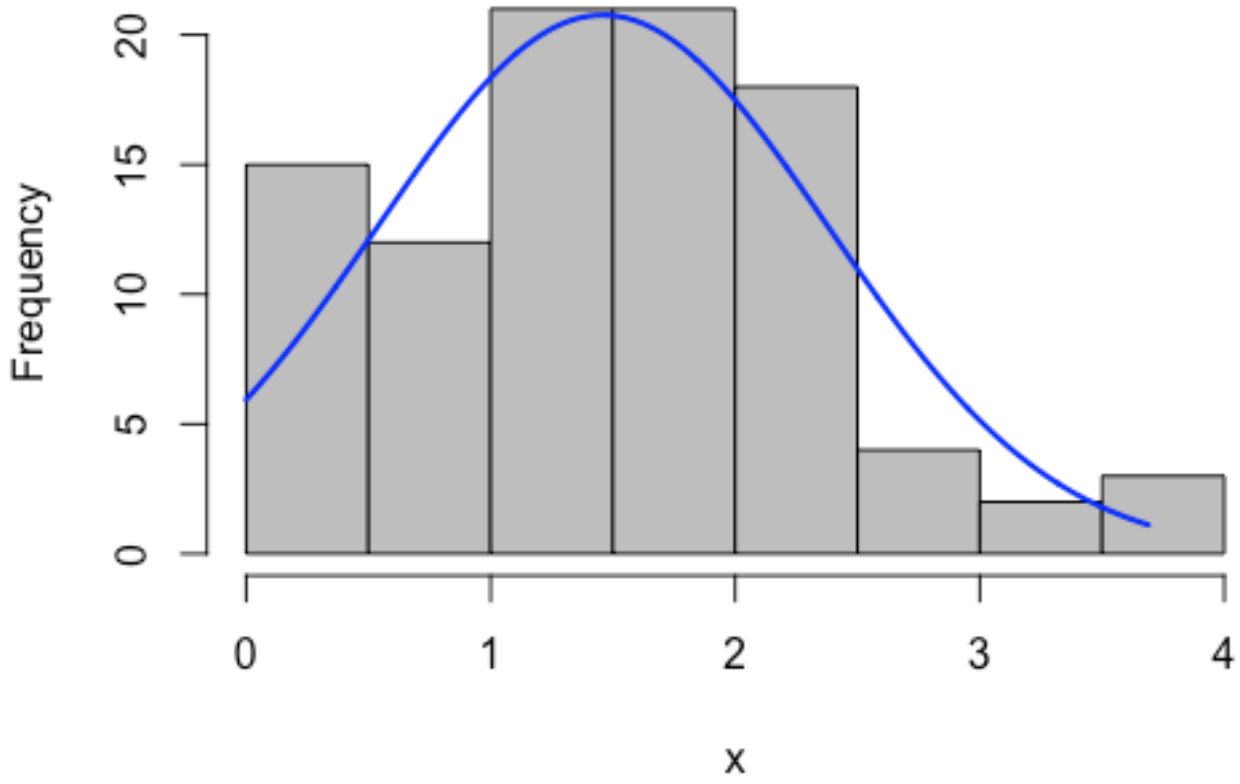

Poisson
0.05
0.10
0.15
0.20
0.25
T = prop_zero_poiss_4var
T(y_rep)
T(y)
Negative Binomial
0.1
0.2
T = prop_zero_4var
T(y_rep)
T(y)


Poisson
Negative Binomial
T = prop_max_poiss_4var
T(y_rep)
T(y)
T = prop_max_4var
T(y_rep)
T(y)
0.025
0.050
0.075
0.100
0.00
0.05
0.10

Comparing log transformed linear model to count model -paradigm shifts-

- Community runs an ANOVA on transformed data
- This is known to be inappropriate for count data
- Does not allow proper accounting for variance
  - Model leaves much on the table which can get assigned to coefficients
- For my data, ANOVA and GLMs tend to find the correct important variables
- Transformation tends to just increase significance

All Monochamus beetles log(x +1) transform



Anova comparisons

Model	Intercept	SE	height low	SE	lures plain	SE
untransformed Anova 4 var	-0.04	1.72	4.83*	1.04	-5.08*	1.04
transformed Anova 2 var	27.2*	0.38	3.96*	0.45	-0.79*	0.45
transformed Anova 4 var	3.12*	0.58	3.96*	0.32	-0.79*	0.32
* significant <.05						

Conclusions

- The best distribution for the zero heavy data is the negative binomial
  - but when all variables are included the model difference is reduced
- Log(X+1) transformation on this data still returns the correct variables in the right direction but struggles with variance in the data
  - Do future tests on other trapping datasets to see the limits of transformation for picking correct predictors
- As expected, adding monochamol significantly raises catch number
- Traps at ground level also increase the catch for these species
- Other significant random effects are the block and date

- datasets-looks good for negbin
  prop_zero <- function(y) mean(y == 0)
  (prop_zero_test1 <- pp_check(stan_model1_nb, plotfun = "stat", stat = "prop_zero"))