Anna Roser

Homework 4: Generative Model for Species Distribution

Study Organism: Lupin

Predictor Variable: Number of pollinators

a. Lupin distribution are likely to be predicted by the number of pollinators in an area because lupin rely on many pollinators for reproduction and fertilization. The plants will have a higher likelihood of reproductive success in dense pollinator areas.
b. My phi value is 0.5 because I would expect that pollinators are likely to prefer some areas over others so there will not be an even distribution of pollinator density across a landscape.
c. The number of points that I can sample is constrained by the number of plots which I can visit. I've chosen to visit 100 plots because lupin plants can be counted quickly.

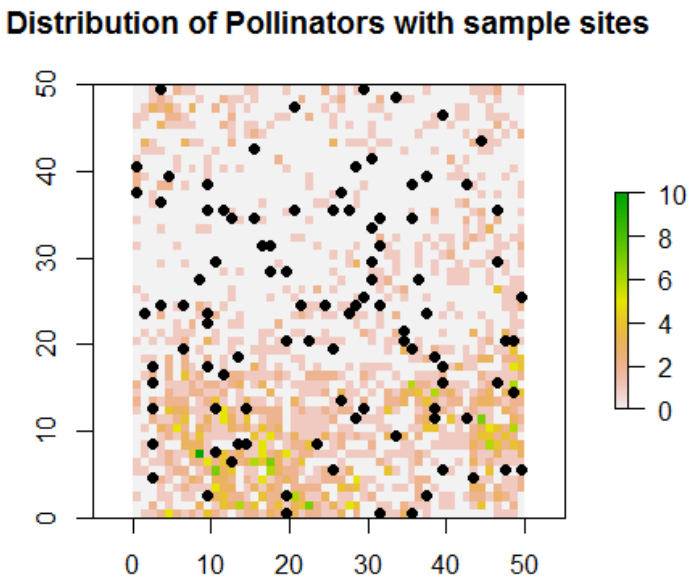


Figure 1. Generated map of pollinator density across an assumed landscape. Sample sites for lupin counts are displayed in black (n=100).

e. We observe a hurdle model because it matches the presence/absence binomial model which are plotted with PA term in code. A hurdle model allows us to better

represent the higher expected number of zeros in presence/absence data; if there are no pollinators present, lupin cannot reproduce and are unlikely to grow in that area. The hurdle model predicts how many 0's we'll get.

f. My landscape has high spatial autocorrelation because pollinators are not evenly distributed, so I would expect that lupin presence is also unevenly distributed over the study site because of plan reliance on pollinators. I assume that the amount of plots with no lupin will be high (0's) but that some plots will have high counts of lupin, resulting in a high variance for the data set.

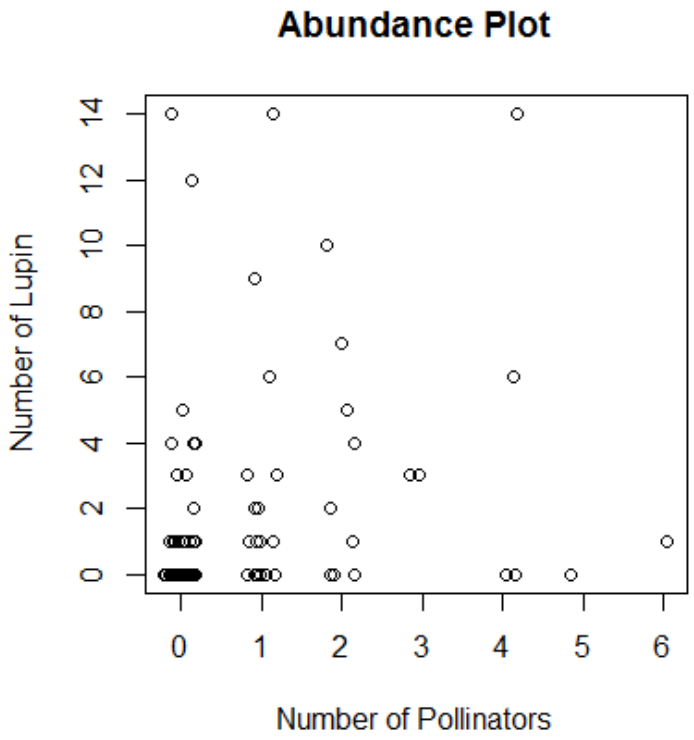


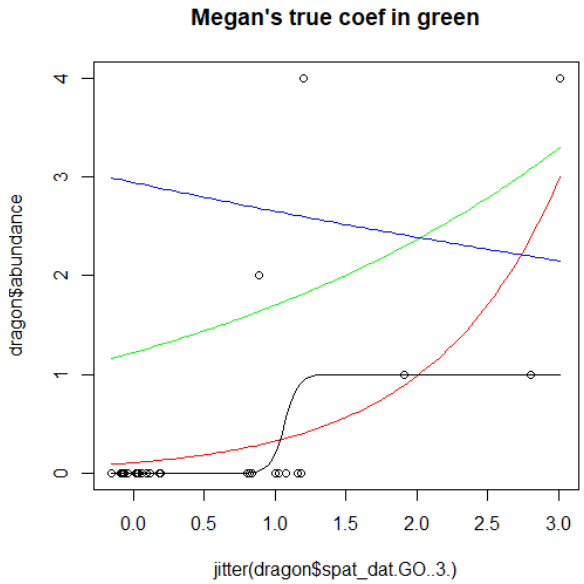
Figure 2. Lupin counts over the number of pollinators at each site.

Neighbor Data: Megan's Dragon's and Castles

I was not able to successful capture the true slope and intercept of Megan's dragons—Is this an issue of small sample size?

True Values:	My Estimates:
#PA int= -3	-21.8

#PA slope= 2.7	20.4
#count int= .2	1.07
#count slope= .33	-0.104



curve(exp(.2 + .33*x), add = T, col = "green") #megan's true coef
curve(exp(-2.226 + 1.104*x), add = T, col = "red") #my estimates of megan's coef
curve(exp(1.08 - 0.1044*x), add = T, col = "blue") #my estimates of megan's coef
curve(plogis(-21.800+20.414*x), add = T, col = "black")

```
> PA1=ifelse(dragon$abundance>0,1,0)  #separate zeros from ones
> glm(PA1~dragon$spat_dat.GO..3., family = "binomial")

Call:  glm(formula = PA1 ~ dragon$spat_dat.GO..3., family = "binomial"
)

Coefficients:
           (Intercept)  dragon$spat_dat.GO..3.
                -21.80                   20.41

Degrees of Freedom: 29 Total (i.e. Null);  28 Residual
Null Deviance:          27.03
Residual Deviance: 10.01    AIC: 14.01

> plogis(-21.800+20.414)-plogis(-21.800)      #difference of PA for dr
agons
[1] 0.2000471    #You are 20% more likely to see a dragon if there's a
castle?
```

```
> abundance2<-dragon$abundance[which(dragon$abundance>0)]  #model only
those greater than zero
> castle<-dragon$spat_dat.GO..3.[which(dragon$abundance>0)]  #make dat
a frames the same length
> glm(abundance2~castle, family= "poisson")   #hurdle model is a good
way to model data if you think drivers can be seperated.

Call:  glm(formula = abundance2 ~ castle, family = "poisson")

Coefficients:
(Intercept)       castle
     1.0800      -0.1044

Degrees of Freedom: 4 Total (i.e. Null);  3 Residual
Null Deviance:           3.942
Residual Deviance: 3.838    AIC: 20.98
> exp(1.0799-0.1044)-exp(1.0799)
[1] -0.291892
```

i. I would plot the autocorrelation of response variable and make a spatial autocorrelation map. If I plotted lupin counts and locations, then I could get an estimate of phi from my lupin density distribution.