

HW 4

Kristina Parker

2/21/2019

Step 1: Study organism

My study organism are rattlesnakes

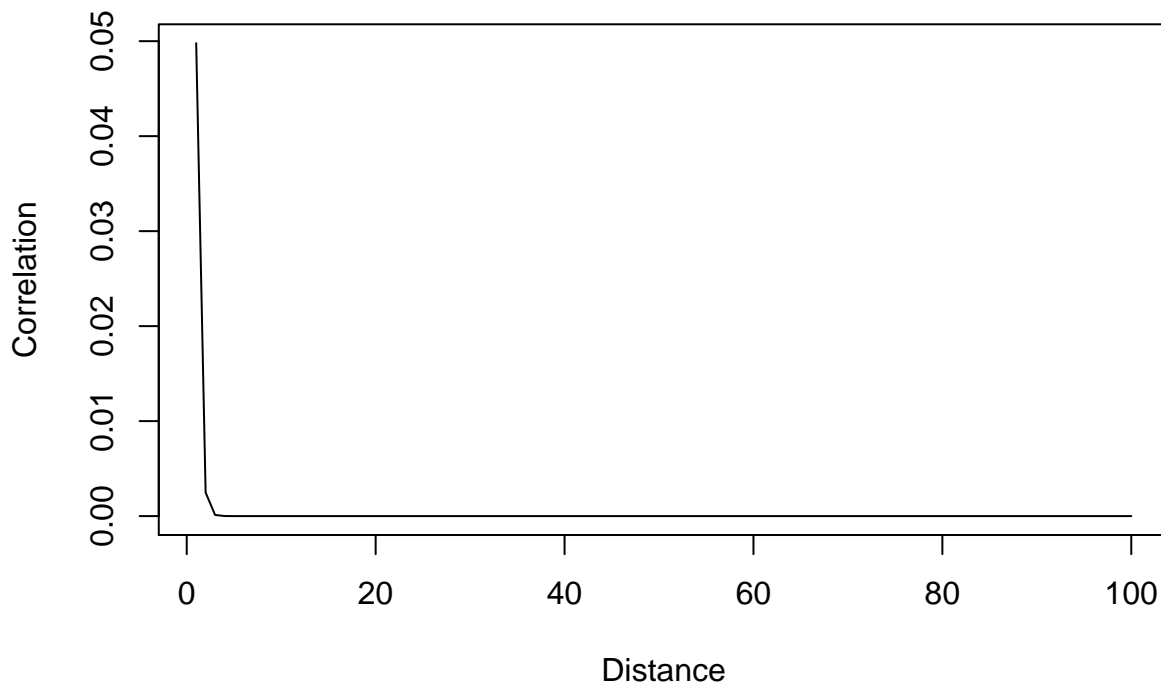
Step 2: Environmental predictor variable that determines abundance of study organism

Number of rock outcroppings in an area

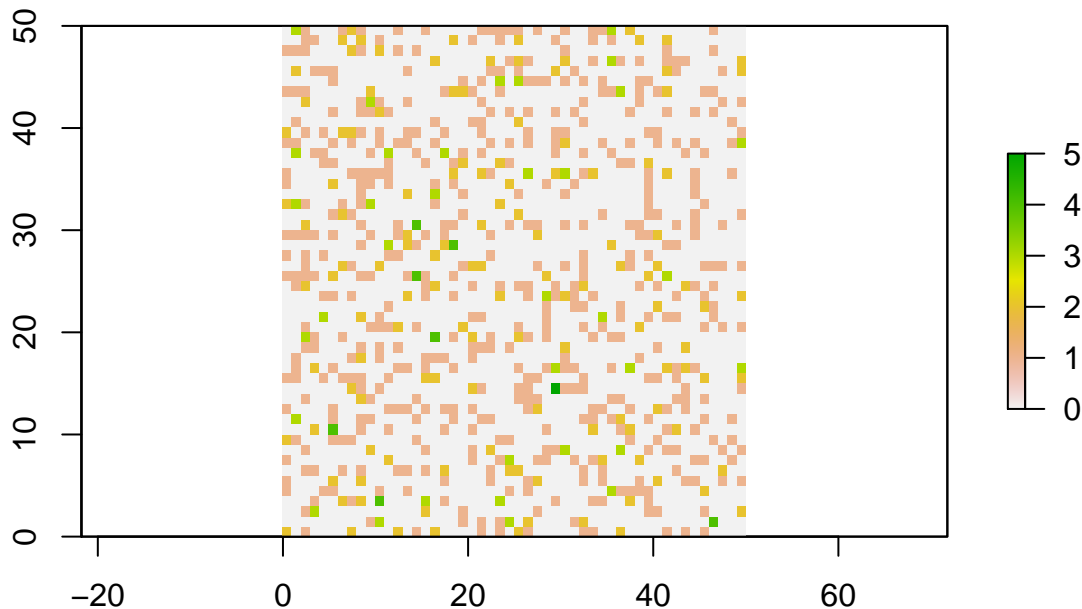
Step 3: Simulate landscape of the environmental variable using Raster function

```
simgrid <- expand.grid(1:50, 1:50)
n <- nrow(simgrid)
distance <- as.matrix(dist(simgrid))

phi = 3
plot(1:100, exp(-phi * 1:100), type = "l", xlab = "Distance", ylab = "Correlation")
```



```
X <- rbinom(n, prob = plogis(-3 + rmvn(1, rep(0,n), exp(-phi * distance))), size = 5)
Xraster <- rasterFromXYZ(cbind(simgrid[, 1:2] - 0.5, X))
plot(Xraster)
```



a)

The number of rock outcroppings in an area is a good predictor variable to represent the distribution of rattlesnakes. Rattlesnakes migrate in the spring from hibernaculums to their summer range which are rock outcroppings. These provide ideal basking locations and nooks to hide from predators. Due to rattlesnakes ambush foraging behavior large rock outcroppings with rodent tunnels are hotspots.

b)

The phi value is 1.7 because I wanted my simulated environment to be less spatially-autocorrelated. This in turn creates a few areas with high number of rock outcroppings while majority of the landscape is vacant.

Step 4: Chose a number of points to sample for your study organism

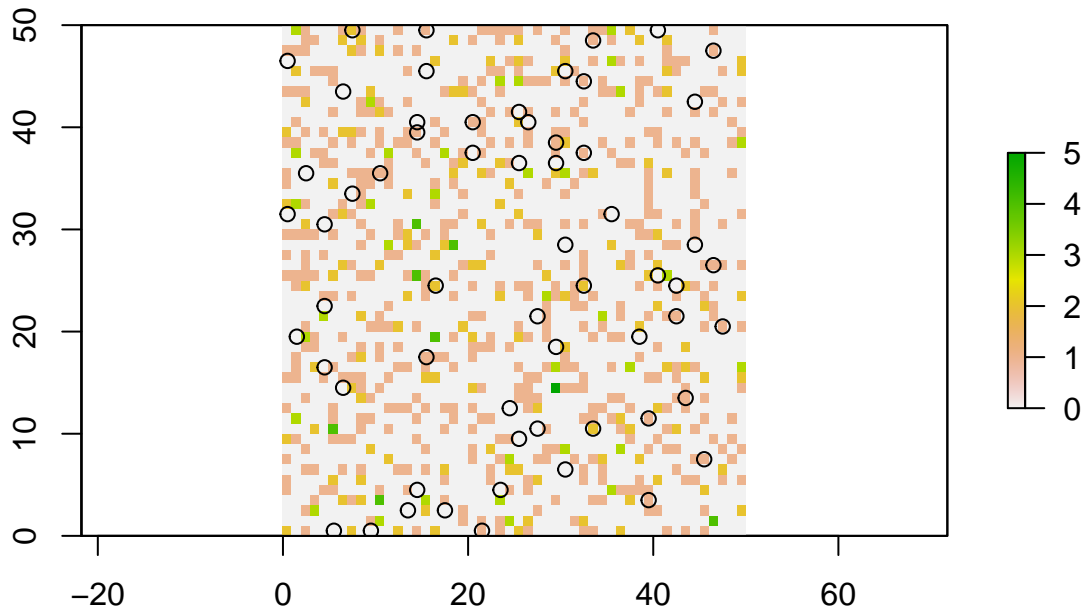
```
spat_dat = rasterToPoints(Xraster)
G0 = sample(x = c(1:nrow(spat_dat)), size = 60)
```

c)

The limiting factor for the number of points that can be sampled is setting the size parameter in the selecting from the raster of the environmental predictor variable. I set mine to 40 samples because of the area size is large; however, there are only a few areas with rock outcroppings. Only the areas searched at a rock outcropping would take time to search, so majority of the sample locations won't take a lot of search effort.

d)

```
plot(Xraster)
points(spat_dat[G0,c(1:2)])
```

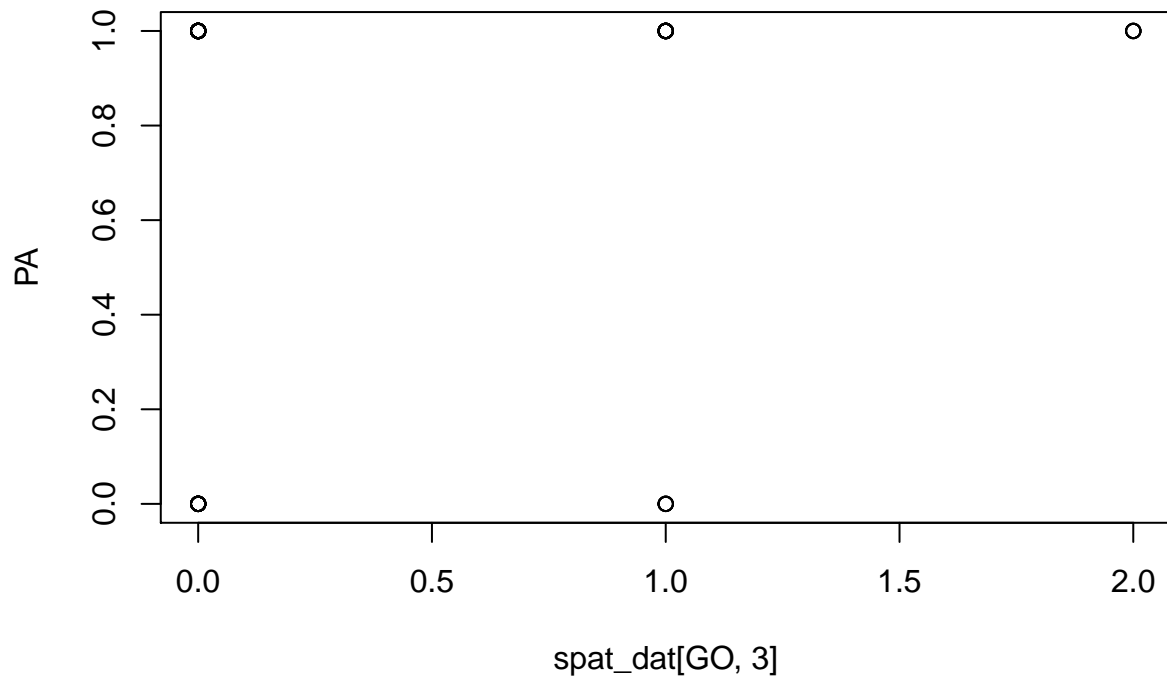


Step 5: Hurdle model/ Zero-inflated negative binomial

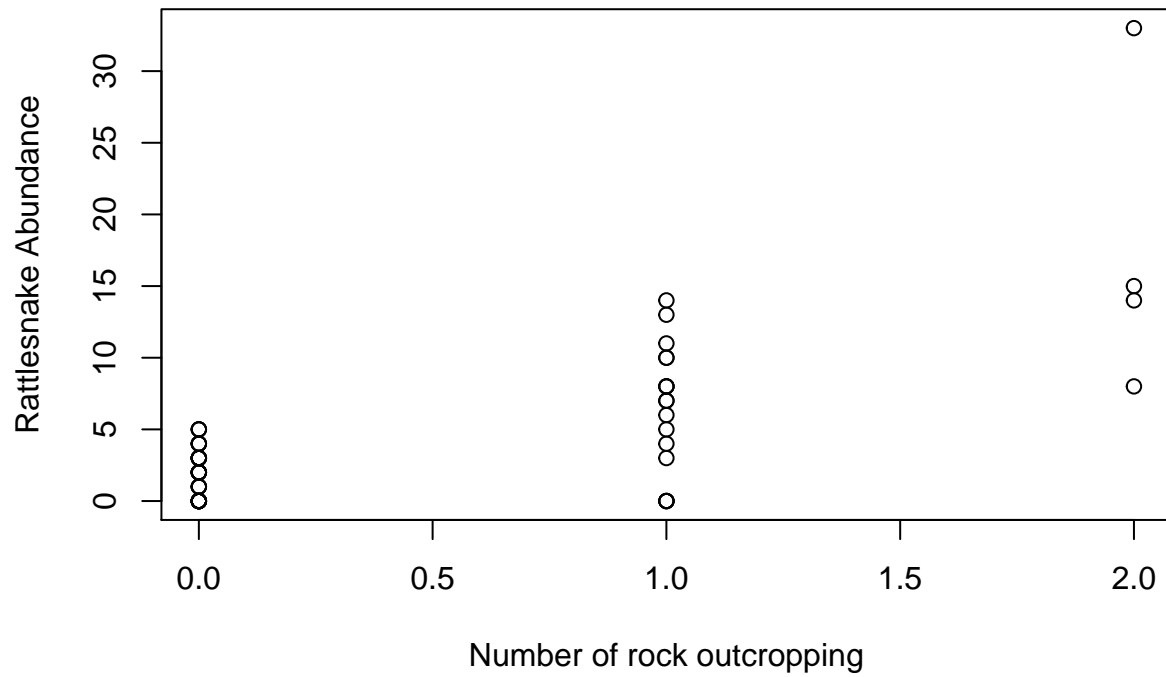
e)

The hurdle model can be used to observe rattlesnakes because they are tightly associated to select habitat types. Using the hurdle models to determine abundance in certain areas is useful because majority of location survey will have low or zero abundance.

f)



```
plot(abundance ~ spat_dat[GO, 3], ylab = "Rattlesnake Abundance", xlab = "Number of rock outcropping")
```



Step 6: Data share

h)

I shared my data with Pat and recieved her data on Northern Goshawks in area of tree canpoy cover is large (>20 cm dbh). I was not able to recapture the true slope and intercept parameters.

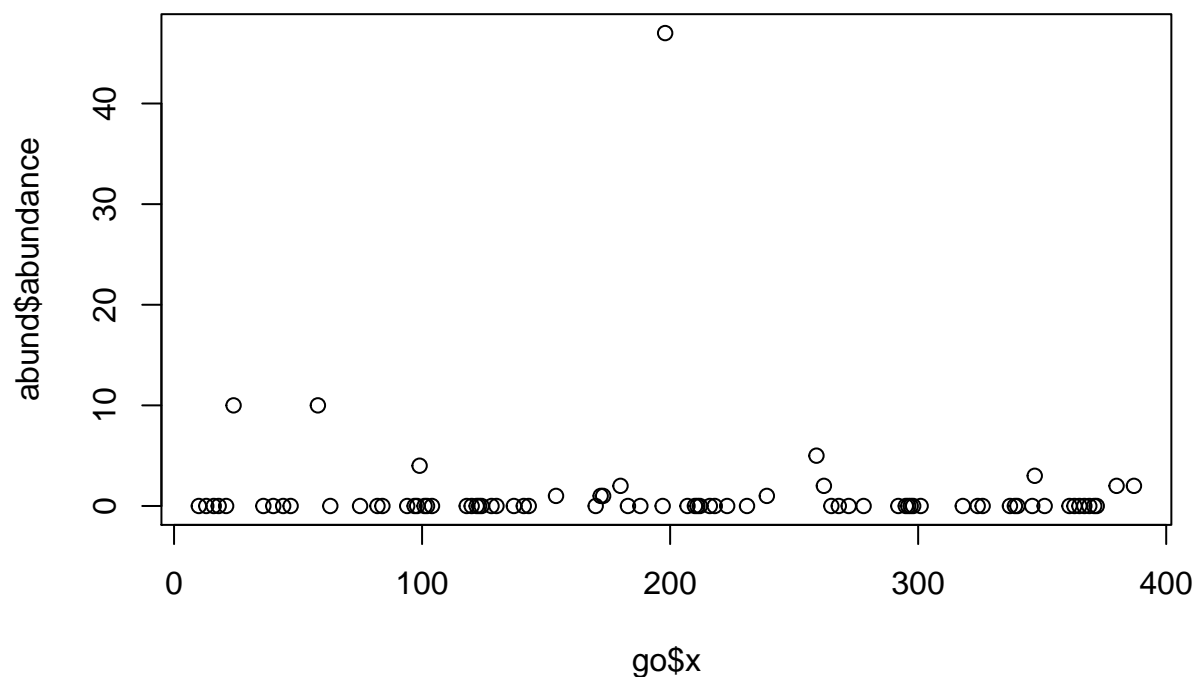
```
abund <- read.csv("/Users/Jojo/Downloads/NOGO_abundance.csv")
str(abund)
```

```
## 'data.frame': 80 obs. of 2 variables:
## $ X : int 1 2 3 4 5 6 7 8 9 10 ...
## $ abundance: int 0 0 0 0 0 0 0 0 1 2 ...
```

```
go <- read.csv("/Users/Jojo/Downloads/NOGO_GO.csv")
str(go)
```

```
## 'data.frame': 80 obs. of 2 variables:
## $ X: int 1 2 3 4 5 6 7 8 9 10 ...
## $ x: int 365 211 296 84 212 337 231 123 173 380 ...
```

```
plot(abund$abundance ~ go$x)
```



```
# PA_True_Slope = -2.5
# PA_True_Slope = 1.5
PA1 = ifelse(abund$abundance > 0, 1, 0)
glm(PA1 ~ go$x, family = "binomial")
```

```
##
## Call: glm(formula = PA1 ~ go$x, family = "binomial")
##
## Coefficients:
## (Intercept)      go$x
## -1.7456316    0.0009595
```

```
##
## Degrees of Freedom: 79 Total (i.e. Null); 78 Residual
## Null Deviance: 74.2
## Residual Deviance: 74.06 AIC: 78.06
# Abund_True_Intercept = 0.5
# Abund_True_Slope = 0.75
abundance2 <- abund$abundance[which(abund$abundance > 0)]
factor <- go$x[which(abund$abundance > 0)]
glm(abundance2 ~ factor, family = "poisson")

##
## Call: glm(formula = abundance2 ~ factor, family = "poisson")
##
## Coefficients:
## (Intercept) factor
## 2.446117 -0.002984
##
## Degrees of Freedom: 13 Total (i.e. Null); 12 Residual
## Null Deviance: 158.2
## Residual Deviance: 149.2 AIC: 195.4
```

i)

To look for spatial-autocorrelation in the response variable I would look at the correlation of the response variable at every sampled location determine if they are spatially correlated.