CalleryHW2Output

Katie Callery

1/23/2018

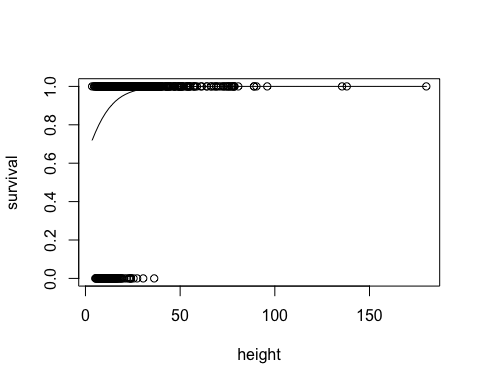
## R Markdown

Question 1

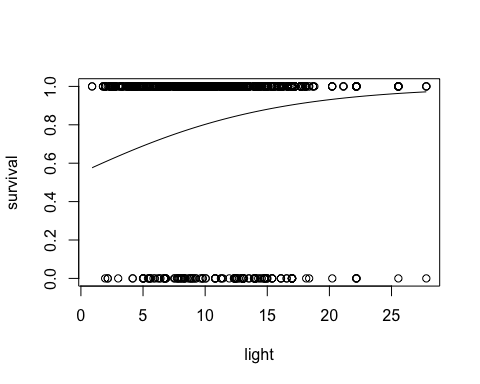
plants<-read.csv("SEEDLING\_SURVIVAL.csv")  
head(plants)

## survival HEIGHT LIGHT  
## 1 1 47.0 2.40  
## 2 1 70.2 14.83  
## 3 1 16.3 9.15  
## 4 1 23.5 8.62  
## 5 1 23.0 4.26  
## 6 1 21.0 3.38

plot(plants$HEIGHT, plants$survival, xlab="height", ylab="survival")  
curve(plogis(0.5+0.125\*x), add=T)



plot(plants$LIGHT, plants$survival, xlab="light", ylab="survival")  
curve(plogis(0.2+0.12\*x), add=T)



heightModel<-glm(plants$survival~plants$HEIGHT, family="binomial")  
coef(heightModel)

## (Intercept) plants$HEIGHT   
## -0.06271111 0.14071141

plogis(-0.06+0.1407\*10)-plogis(-0.06+0.1407\*5)

## [1] 0.1380945

# 0.138 is the difference between 5 and 10.   
0.138/5

## [1] 0.0276

# 0.0276 is the between 5 and 10.  
maxSlopeHeight=0.1407/4  
#0.035 is the maximum slope.

lightModel<-glm(plants$survival~plants$LIGHT, family="binomial")  
coef(lightModel)

## (Intercept) plants$LIGHT   
## 2.66194692 -0.06552684

plogis(2.6-0.066\*10)-(2.6-0.066\*5)

## [1] -1.395648

# -1.395648 is the difference between 5 and 10  
-1.395648/5

## [1] -0.2791296

#-0.2791296 is the slope between 5 and 10  
maxSlopeLight<-(-0.066/4)  
#-0.0165 is the maximum slope.

confint(heightModel)

## Waiting for profiling to be done...

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## 2.5 % 97.5 %  
## (Intercept) -0.5791061 0.4268167  
## plants$HEIGHT 0.1038803 0.1815477

confint(lightModel)

## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) 2.25136434 3.0876309  
## plants$LIGHT -0.09841747 -0.0325795

Height is a better predictor of seedling survival than light. Light seem to have a significant negative effect on survival based on the slope parameter of the logistic model and the confidence intervals of the linear equation not overlapping with zero. Height seems to have a significant positive effect on survival based on the slope parameter of the logistic model and the confidence intervals of the linear equation not overlapping with zero.

QUESTIONS: *How do you get the confidence intervals of the logistic model? They would be changing as the slope changes, right? Is it not necessary to calculate them?*

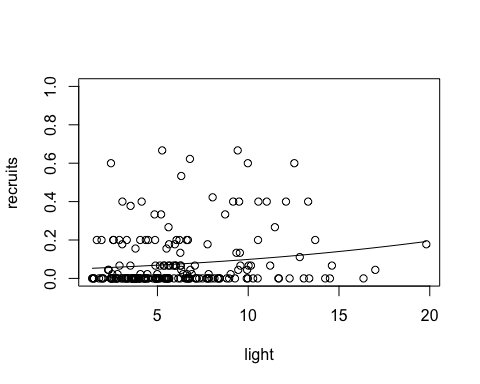
Question 2

seeds<-read.csv("Seeds.csv")  
head(seeds)

## Site Pile DBH seedlings seeds recruits grass light  
## 1 m1 m1.15 21.590 0 15 2 1 9.35  
## 2 m1 m1.45 0.000 0 45 2 0 17.00  
## 3 m1 m1.5 46.990 0 5 1 0 6.68  
## 4 m10 m10.15 0.000 0 15 0 0 6.72  
## 5 m10 m10.45 27.686 0 45 0 0 4.91  
## 6 m10 m10.5 0.000 0 5 2 0 3.07

1. I chose light as the predictor variable for recruitment

seeds$probSuccess<-seeds$recruits/seeds$seeds  
plot(seeds$light, seeds$probSuccess, xlab="light", ylab="recruits")  
curve(plogis(0.079\*x-3),add=T)

 b. create a proportional variable for the binomial glm

response<-cbind(seeds$recruits, seeds$seeds-seeds$recruits)  
lightRecruits<-glm(response~seeds$light, family="binomial")  
coef(lightRecruits)

## (Intercept) seeds$light   
## -3.0936296 0.0798368

#Maximum slope:  
0.079/4

## [1] 0.01975

# Maximum slope is 0.01975.  
# Slope between 2 points  
plogis(-3+0.079\*10)-plogis(-3+0.079\*5)

## [1] 0.03003875

#0.03003875 is the difference between 10 and 5.   
0.03003875/5

## [1] 0.00600775

#0.00600775 is the slope between 10 and 5

1. confidence intervals for slope and intercept

confint(lightRecruits)

## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) -3.32799150 -2.8631191  
## seeds$light 0.04990323 0.1088193

Light has a small positive effect on survival, and it is significant because the confidence intervals do not overlap zero.

Question 3

a.) b.)

mosquito<-read.csv("mosquito\_data.csv")  
head(mosquito)

## Emergent\_adults Egg\_Count Detritus  
## 1 2 3 0.00  
## 2 0 2 0.01  
## 3 2 3 0.01  
## 4 4 6 0.02  
## 5 1 4 0.02  
## 6 5 7 0.03

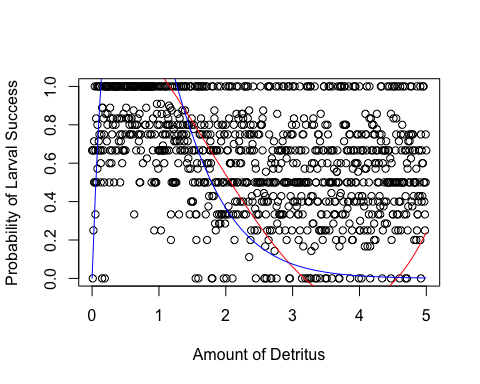
attach(mosquito)  
mosquito$success<-Emergent\_adults/Egg\_Count  
attach(mosquito)

## The following objects are masked from mosquito (pos = 3):  
##   
## Detritus, Egg\_Count, Emergent\_adults

plot(Detritus, success, xlab="Amount of Detritus", ylab="Probability of Larval Success")  
larvalSuccess<-cbind(Emergent\_adults, Egg\_Count-Emergent\_adults)  
glm(larvalSuccess~Detritus, family="binomial")

##   
## Call: glm(formula = larvalSuccess ~ Detritus, family = "binomial")  
##   
## Coefficients:  
## (Intercept) Detritus   
## 1.3240 -0.3216   
##   
## Degrees of Freedom: 999 Total (i.e. Null); 998 Residual  
## Null Deviance: 1564   
## Residual Deviance: 1331 AIC: 2881

#(Intercept) Detritus   
# 1.3240 -0.3216   
  
curve(1.44 - 0.19\*x - 0.21\*x^2 + 0.04\*x^3, add=T, col="red") #polynomial curve  
curve(10\*x\*exp(-2\*x), add=T, col="blue") #Ricker curve

 c.) The biological implication that is different between the models is that the Polynomial model predicts an increase in larval success when detritus is greater than 4, while the Ricker model does not predict such an increase.

d.) Find maximum likelihood estimate through minimum negative log-likehood comparisons

# for the polynomial  
-sum(dbinom(x = Emergent\_adults, size=Egg\_Count, prob=plogis(1.44 - 0.19\*Detritus - 0.21\*Detritus^2 + 0.04\*Detritus^3), log=T))

## [1] 1415.63

# for the Ricker  
-sum(dbinom(x = Emergent\_adults, size=Egg\_Count, prob=plogis(10\*Detritus\*exp(-2\*Detritus)), log=T))

## [1] 1385.847

#Polynomial = 1415.63  
#Ricker = 1385.847  
  
?dbinom  
  
# What does plogis do?  
# What does dbinom do?  
# How would you plot the maximum likelihood curves for this?

e.) The likelihood is higher for the Ricker model.

Question 4

# Example for-loop  
  
empty\_vector<-rep(NA,times=100)  
for(i in 1:100){ empty\_vector[i]<-rnorm(n=1,mean=0,sd=1)}

Reed et al. 2013, Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine (*Parus major*).

Number of nestlings produced ~ individual (female) breeding mismatch with prey peak

# variables  
  
slope<--0.015  
intercept<-2.555  
stdev<-0.2  
  
# predictor variable  
  
#mismatchDays<-round(runif(100, min=-60, max=60))  
  
# create a vector of samples sizes  
  
sample\_size<-c(100:5000)  
  
# create a vector to fill with results of the power analysis the same length as the sample size vector  
  
length(sample\_size) # = 4901

## [1] 4901

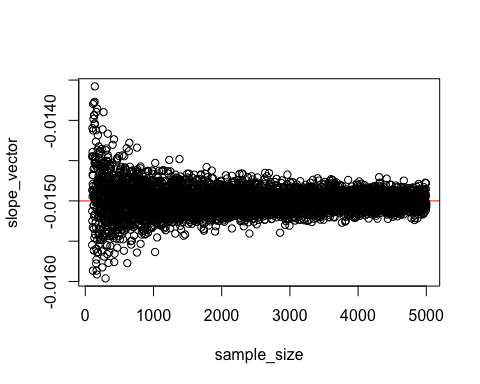
# rep is repeat, NA is empty, and times is 4901  
slope\_vector<-rep(NA,times=4901)  
power\_vector<-rep(NA,times=4901)

Create the for loop!

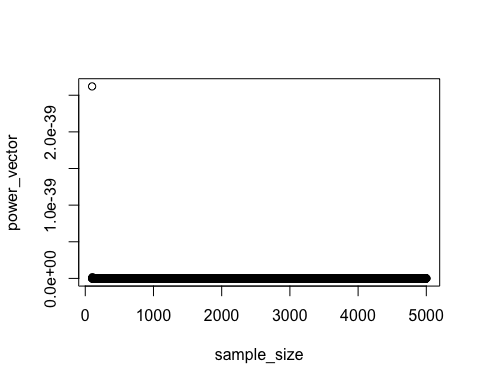
# For loop that will do the power analysis:  
# Each iteration will create a random normal sample of the data at the size specified by the iteration. Then it will create a model of that random data against the predictor variable. Then I will extract the slope estimate from the model, and store it in that iteration's cell in the power vector  
  
for(i in 1:4901){  
  
mismatchDays<-round(runif(i+99, min=-60, max=60))  
  
numFledge<-rnorm(n=sample\_size[i], mean=(intercept+slope\*mismatchDays), sd=stdev)  
  
FledgeMod<-glm(numFledge~mismatchDays)  
  
SumFledge<-summary(FledgeMod)  
  
getCoef<-coefficients(SumFledge)  
  
getSlope<-getCoef[2,1]  
getPower<-getCoef[2,4]  
  
slope\_vector[i]<-getSlope  
  
power\_vector[i]<-getPower  
  
}

Plot the output. The slope or power is plotted on the y-axis, and the sample size (100:5000) is plotted on the x-axis.

plot(sample\_size, slope\_vector, abline(h=-0.015,col="red"))



plot(sample\_size, power\_vector, abline(h=0.05, col="red", xlim=c(0,10)))

 I’m not sure why my power graph looks like this… It should be centering around -0.015…. Am I doing something wrong with simulating my data?

# Binomial data - I made this up.   
  
slope\_vecb<-rep(NA,times=50)  
power\_vecb<-rep(NA,times=50)  
  
sample\_size<-c(1:50)  
  
for(i in 1:50){  
  
predictor<-runif(sample\_size[i], min=0, max=1)  
  
response<-rbinom(n=sample\_size[i], size=1, prob=plogis(0.001+0.0025\*predictor))  
  
model<-glm(response~predictor, family = "binomial")  
  
modelSum<-summary(model)  
  
coefMod<-coef(modelSum)  
  
slopeb<-coefMod[2]  
  
slope\_vecb[i]<-slopeb  
  
}

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

plot(sample\_size, slope\_vecb, ylim=c(0,1))

