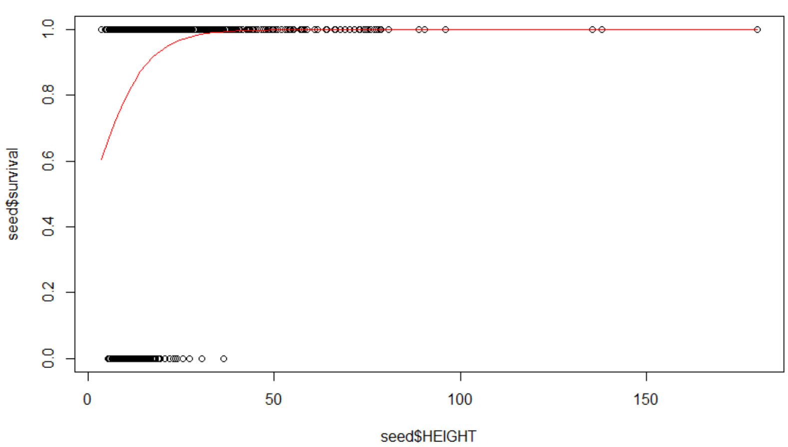
Megan Kelly-Slatten
2-5-19

Homework 2

1. **Seedling survival based on Height and Light**

Height

```
> seed<-read.csv("L:/Home/S243/Lab Members/Megan Kelly Slatten/SEEDLING_SURVI
VAL.csv")
> summary(seed)
    survival          HEIGHT           LIGHT      
 Min.   :0.0000   Min.   :  3.50   Min.   : 0.90  
 1st Qu.:1.0000   1st Qu.: 11.00   1st Qu.: 7.64  
 Median :1.0000   Median : 13.80   Median :10.33  
 Mean   :0.8725   Mean   : 18.33   Mean   :10.76  
 3rd Qu.:1.0000   3rd Qu.: 19.80   3rd Qu.:13.38  
 Max.   :1.0000   Max.   :180.00   Max.   :27.80  
> str(seed)
'data.frame':	1435 obs. of  3 variables:
 $ survival: int  1 1 1 1 1 1 1 1 1 1 ...
 $ HEIGHT  : num  47 70.2 16.3 23.5 23 21 30 17.5 76 57.5 ...
 $ LIGHT   : num  2.4 14.83 9.15 8.62 4.26 ...
> plot(seed$survival~seed$HEIGHT)
> m1<-glm(seed$survival~seed$HEIGHT, family="binomial")
> coef(m1)
(Intercept) seed$HEIGHT 
-0.06271111  0.14071141 
> plogis(-0.06271111)
[1] 0.4843274
> 0.1407/4
[1] 0.035175
> curve(plogis(-0.062711+0.140711*x), add=T, col="red")
```

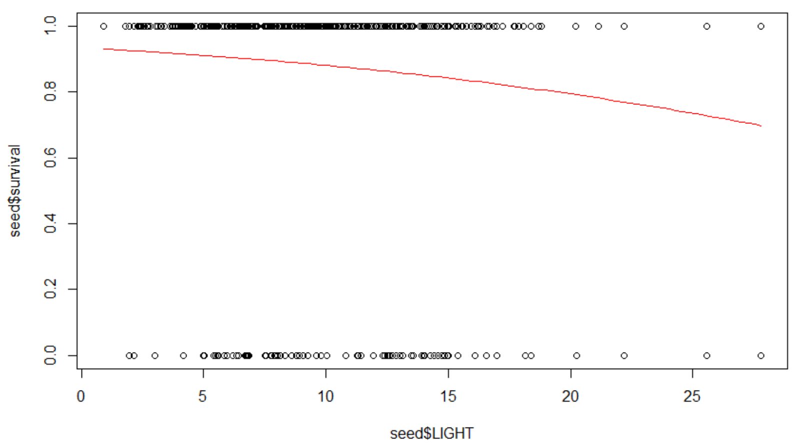



```
> confint(m1)
                 2.5 %    97.5 %
(Intercept) -0.5791061 0.4268167
seed$HEIGHT  0.1038803 0.1815477
> plogis(-0.062711+0.140711*50)-plogis(-0.062711+0.140711*5)
[1] 0.3441197
```

When seed height is zero, 48% of seeds survive. The maximum increase in seedling survival as you increase by one height unit is 3.5%. The increase from a 5cm seedling to a 50cm increases seedling survival by 34.4%.

Light

```
> plot(seed$survival~seed$LIGHT)
> m2<-glm(seed$survival~seed$LIGHT, family="binomial")
> coef(m2)
(Intercept)  seed$LIGHT
 2.66194692 -0.06552684
> plogis(2.66194692 )
[1] 0.9347435
> -0.0655/4
[1] -0.016375
> curve(plogis(2.66194+-0.065526*x), add=T, col="red")
```



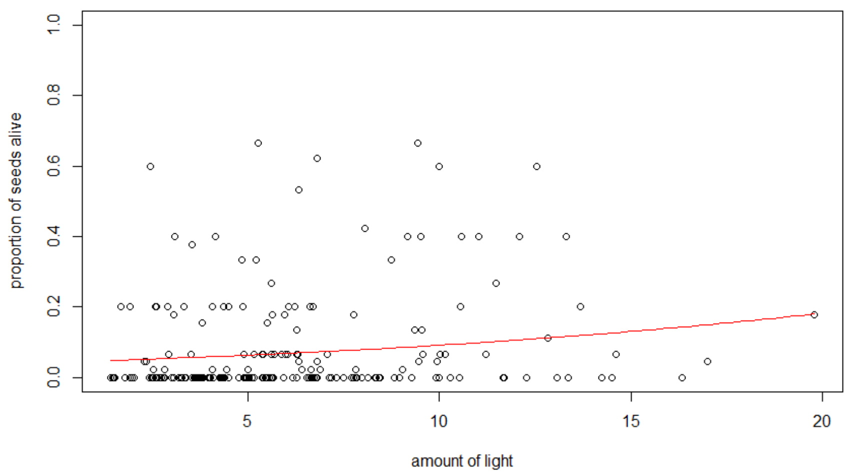

```
> confint(m2)
                  2.5 %     97.5 %
(Intercept)  2.25136434  3.0876309
seed$LIGHT  -0.09841747 -0.0325795
> plogis(2.66194+-0.065526*25)-plogis(2.66194+-0.065526*5)
[1] -0.1759688
```

When there is no light, seedling survival is 93%. The maximum effect of seedling survival as you increase by one unit of light is -1.6%. The increase in light from 5 to 25 has a -18% decrease on seedling survival. Height is a stronger predictor of seedling survival.

2. **Light effects on Seedling Success**

```
> plant<-read.csv("L:/Home/S243/Lab Members/Megan Kelly Slatten/School/mod st
ats/Seeds.csv")
> str(plant)
'data.frame':   281 obs. of  8 variables:
 $ Site      : Factor w/ 94 levels "m1","m10","m11",..: 1 1 1 2 2 2 3 3 3 4 ..
.
 $ Pile      : Factor w/ 281 levels "m1.15","m1.45",..: 1 2 3 4 5 6 7 8 9 10 .
..
 $ DBH       : num  21.6 0 47 0 27.7 ...
 $ seedlings: int  0 0 0 0 0 0 0 0 0 0 ...
 $ seeds     : int  15 45 5 15 45 5 15 45 5 15 ...
 $ recruits : int  2 2 1 0 0 2 6 0 1 1 ...
 $ grass     : int  1 0 0 0 0 0 0 1 1 0 ...
 $ light     : num  9.35 17 6.68 6.72 4.91 ...
```

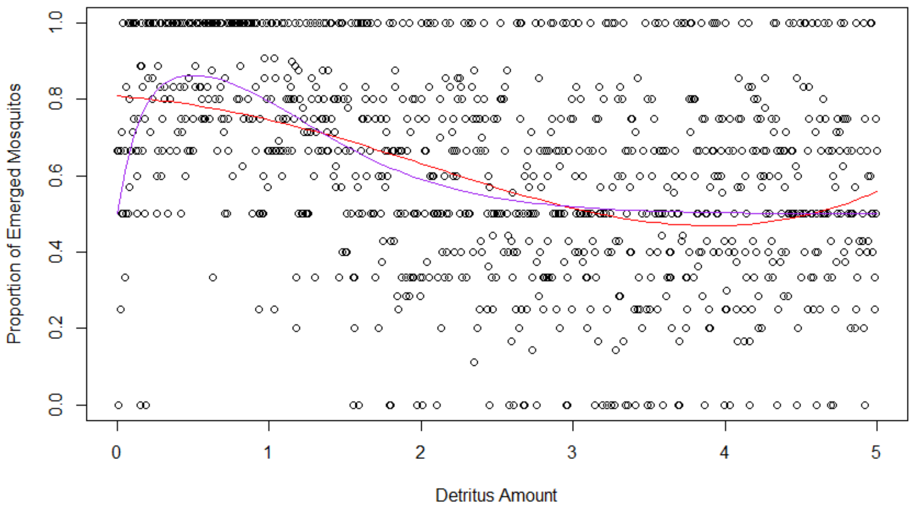
```
> propseed<-(plant$recruits/plant$seeds)
> seedlife<-cbind(plant$recruits, plant$seeds-plant$recruits)
> plot(propseed~plant$light, xlab= "amount of light", ylab="proportion of see
ds alive")
> lightmod<-glm(seedlife~plant$light, family="binomial")
> coef(lightmod)
(Intercept) plant$light
 -3.0936296   0.0798368
> plogis( -3.0936296)
[1] 0.04337079
> 0.0798368/4
[1] 0.0199592
> confint(lightmod)
                  2.5 %     97.5 %
(Intercept) -3.32799150 -2.8631191
plant$light  0.04990323  0.1088193
> curve(plogis(-3.0936+0.0798*x), add=T, col="red")
```



When there is no light, 4% of the seeds germinate. As you increase light by one unit, the maximum increase in the proportion of seeds that germinate is by 2%. The 95% CI slope does not overlap zero, and so I conclude that the amount of light has an effect on seedling survival.

3. Mosquitos

```
> buggy<-read.csv("L:/Home/S243/Lab Members/Megan Kelly Slatten/School/mod st
ats/mosquito_data.csv")
> str(buggy)
'data.frame':   1000 obs. of  3 variables:
 $ Emergent_adults: int  2 0 2 4 1 5 2 1 1 4 ...
 $ Egg_Count      : int  3 2 3 6 4 7 4 1 2 6 ...
 $ Detritus       : num  0 0.01 0.01 0.02 0.02 0.03 0.03 0.04 0.04 0.05 ...
> hatched<-(buggy$Emergent_adults/buggy$Egg_Count)
> plot(hatched~buggy$Detritus, ylab= "Proportion of Emerged Mosquitos", xlab=
"Detritus Amount")
> curve(plogis(1.44-0.19*x-0.21*x^(2)+0.04*x^(3)),add=T, col="red")
> curve(plogis(10*x*(exp(-2*x))), add=T, col="purple")
```



The polynomial model has a gradual change in slope as you move from low to high detritus levels, indicating that detritus concentration has an effect on mosquito emergence but that it is slow and gradual. Whereas the Ricker model has a steeper slope that indicates that Detritus concentrations greater than 1 can have dramatic negative impact on mosquito emergence.

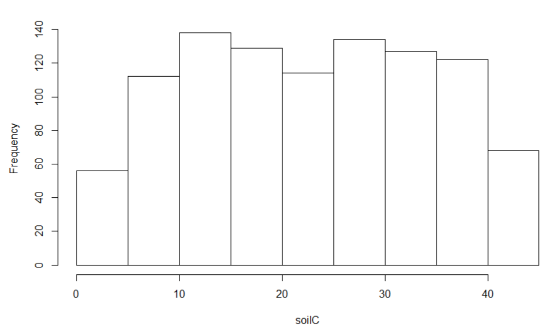
```
-sum(dbinom(x=buggy$Emergent_adults, size=1000, prob=plogis(1.44-0.19*buggy$D
etritus-0.21*buggy$Detritus^(2)+0.04*buggy$Detritus^(3)), log=T))
[1] 1415.63
>
> -sum(dbinom(x=buggy$Emergent_adults, size=1000, prob=plogis(10*buggy$Detrit
us*(exp(-2*buggy$Detritus))),log=T))
[1] 1385.847
```

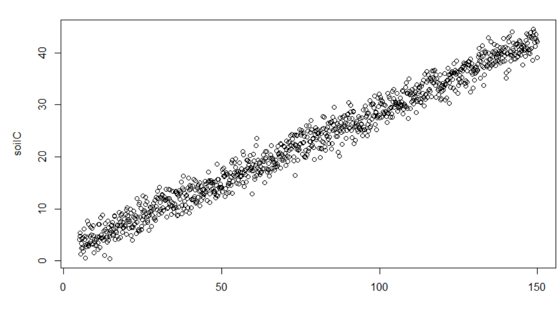
The Ricker model has the lower value, and thus fits the data better.

4. POWER

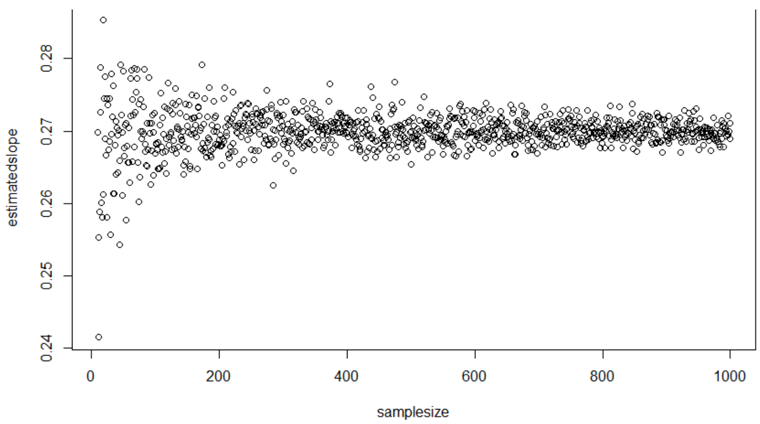
Linear: Mass of plant predicts amount of soil carbon.

```
> intercept=2
> slope=0.27
> n=1000
> mass=seq(from=5, to=150, length=1000)
> soilC=rnorm(n=1000, mean=intercept+slope*mass, sd=1.7)
> hist(soilC)
> plot(soilC~mass)
```



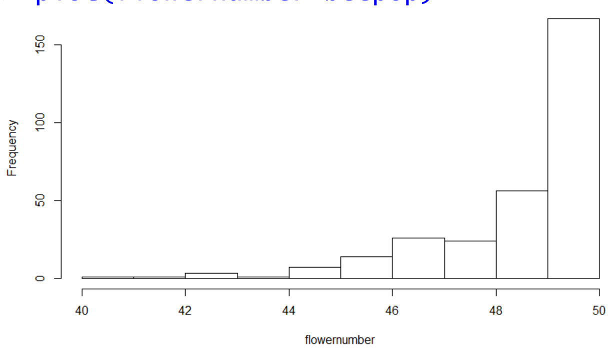


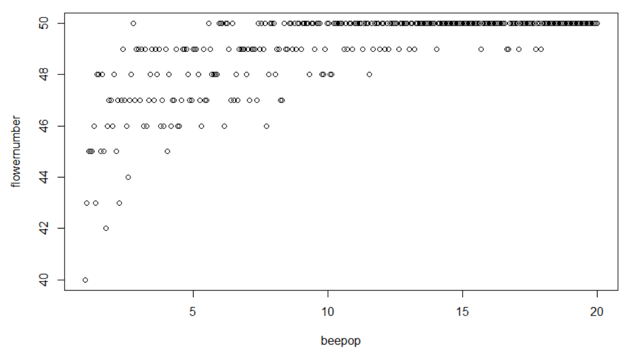
```
> samplesize=seq(from=10, to=1000)
> estimatedslope=rep(NA, times=length(samplesize))
> for(m in 1:length(samplesize)) {
+   y=rnorm(n=samplesize[m], mean=intercept+slope*seq(from=5, to=150, length=
samplesize[m]), sd=1.7)
+   response=y
+   mod1<-glm(response~seq(from=5, to=150, length=samplesize[m]))
+   estimatedslope[m]=coef(mod1)[2]
+ }
> plot(estimatedslope~samplesize)
```



Binomial: Number of flowering plants/all plants in plot

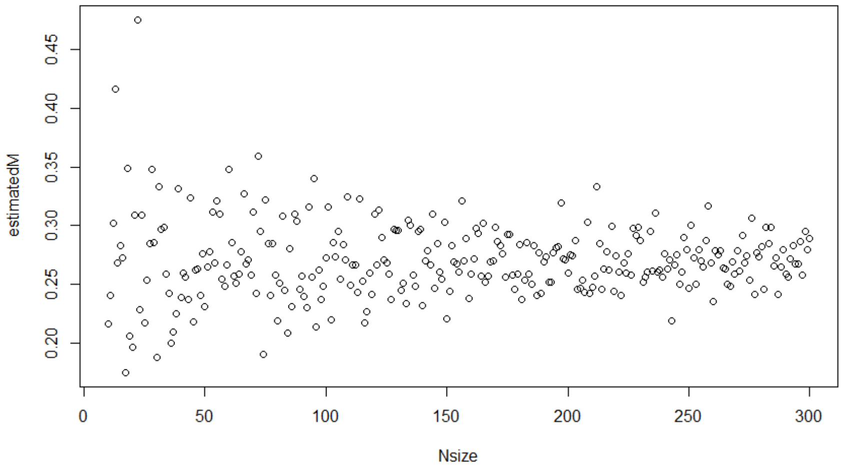
```
> b=2
> m=0.27
> n=300
> beepop=seq(from=1, to=20, length=300)
> flowernumber=rbinom(n=300, prob=plogis(b+m*beepop), size=50)
> hist(flowernumber)
> plot(flowernumber~beepop)
```






```
> Nsize=seq(from=10, to=300)
> estimatedM=rep(NA, times=length(Nsize))
> for(z in 1:length(Nsize)) {
+   y=rbinom(n=Nsize[z], prob=plogis(b+m*seq(from=1, to=20, length=Nsize[z]))
, size=50)
+   response=cbind(y, 50-y)
+   flowermod<-glm(response~seq(from=1, to=20, length=Nsize[z]), family="bino
mial")
+   estimatedM[z]=coef(flowermod)[2]}
```

```
> plot(estimatedM~Nsize)
```



For my linear regression, all of my samples sizes were close to my true slope of 0.27, but around a sample size of 200 I would get the most precise results. For my binomial data, there was a larger spread around my slope. Even with 300 samples, I still had quite a bit of variation around my slope. With discrete data like binomial data, you have larger gaps or jumps between samples and this can lead to more variation, whereas with continuous data the gaps between samples are typically, smaller and therefore there is normally less variation or noise.