PATRICIA KAYE T. DUMANDAN   **03 FEB 2019**

MS RAPTOR BIOLOGY   **EEB 622 HW 2**

Q1: SEEDLING SURVIVAL

> str(q1)

'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 ...

A. HEIGHT

1. model

> q1.glm=glm(q1$survival~q1$HEIGHT, family="binomial")
> coef(q1.glm)

(Intercept)   q1$HEIGHT
-0.06271111 0.14071141

> summary(q1.glm)

Call:

glm(formula = q1$survival ~ q1$HEIGHT, family = "binomial")

Deviance Residuals:

   Min   1Q Median   3Q   Max
-3.1785 0.1743 0.4765 0.6071 1.0012

Coefficients:

        Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.06271   0.25674 -0.244   0.807
q1$HEIGHT   0.14071   0.01984   7.093 1.31e-12 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

   Null deviance: 1095.35 on 1434 degrees of freedom
Residual deviance: 997.64 on 1433 degrees of freedom
AIC: 1001.6

Number of Fisher Scoring iterations: 7

2. Confidence Intervals

> confint(q1.glm)

        2.5 %   97.5 %
(Intercept) -0.5791061 0.4268167
q1$HEIGHT   0.1038803 0.1815477

3.Graph

> plot(q1$survival~q1$HEIGHT, ylab="seedling survival", xlab="height")
> curve(plogis(-0.06271111+0.14071141*x), add=T, col="red", lwd=3)
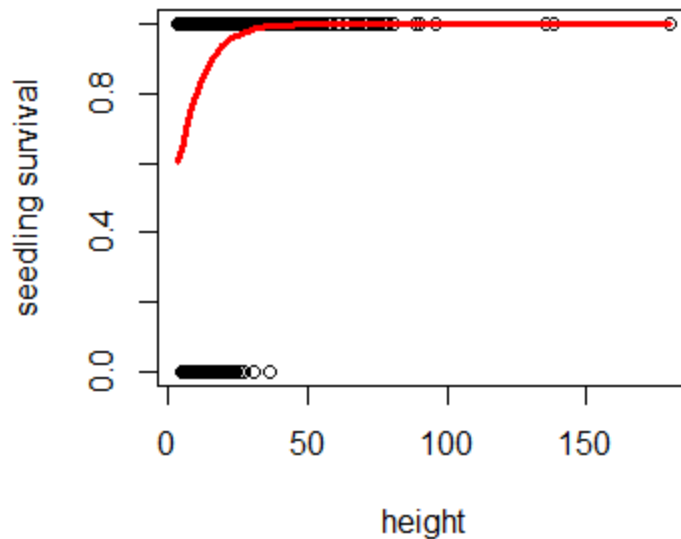
Fig.1. Effect of tree height on seedling survival. Probability of successful seedling growth (survival) is high when tree height is <100m but > ~30m. Seedling survival is at -0.06 probability if tree height is 0. For every 1 unit increase in tree height, survival increases by 0.14 (P=1.31e-12). Predictor (height) resulted to 97.71 points in deviance. decrease Red line shows line of best (AIC=1001.6).

B. **LIGHT**

1. model

```
> q1.1glm=glm(q1$survival~q1$LIGHT, family="binomial")
> coef(q1.1glm)
(Intercept)   q1$LIGHT
 2.66194692 -0.06552684
> summary(q1.1glm)
Call:
glm(formula = q1$survival ~ q1$LIGHT, family = "binomial")
Deviance Residuals:
   Min    1Q   Median    3Q    Max
-2.2849  0.4376  0.4908  0.5512  0.8471
Coefficients:
        Estimate Std. Error z value Pr(>|z|)
(Intercept) 2.66195   0.21315 12.488  < 2e-16 ***
q1$LIGHT   -0.06553   0.01677 -3.906 9.37e-05 ***
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
   Null deviance: 1095.4  on 1434  degrees of freedom
Residual deviance: 1080.3  on 1433  degrees of freedom
AIC: 1084.3
Number of Fisher Scoring iterations: 4
```
2.confidence intervals
```
> confint(q1.1glm)
```

```
          2.5 %    97.5 %
(Intercept)  2.25136434  3.0876309
q1$LIGHT    -0.09841747 -0.0325795
```

3.graph

```
> plot(q1$survival~q1$LIGHT, ylab="seedling survival", xlab="light")
> curve(plogis(2.66194692+-0.06552684*x), add=T, col="red", lwd=3)
```
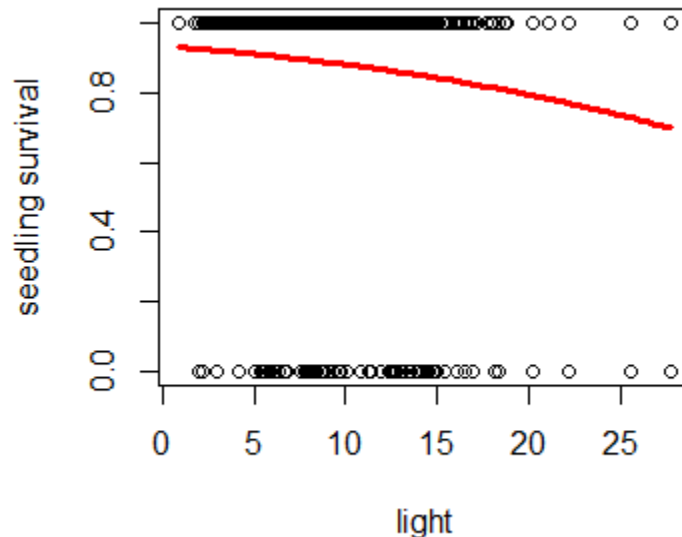

Fig.2. Effect of light on seedling survival Probability of successful seedling growth (survival) is high when light is <20 units but > 0 units. Seedling survival is at 2.66 probability if light is 0. For every 1 unit increase in light, survival decreases by 0.07 (P=9.37e-05). Predictor (light) resulted to 15.1 decrease in deviance. Red line shows line of best (AIC=1080.3).

- **If we compare the AIC of the two models (height and light as predictors), height is a better predictor of seedling survival.**

Q2: SEEDS DATA

1. TEST ASSUMPTIONS

```
> #normality
> shapiro.test(q2$recruits)
        Shapiro-Wilk normality test
data:  q2$recruits
W = 0.5393, p-value < 2.2e-16
> fligner.test(q2$recruits, q2$seedlings)
        Fligner-Killeen test of homogeneity of variances
data:  q2$recruits and q2$seedlings
Fligner-Killeen:med chi-squared = 32.039, df = 6, p-value = 1.604e-05
> #data non-normal
```

2. PROPORTIONAL DATA (SUCCESS)

```
> #make proportions
> prop2=(q2$recruits/q2$seeds)
> rec=cbind(q2$recruits, q2$seeds-q2$recruits)
```

```
> head(rec)
   [,1] [,2]
[1,]  2  13
[2,]  2  43
[3,]  1   4
[4,]  0  15
[5,]  0  45
[6,]  2   3
```

3. PROPORTIONAL GLM

```
> q2.dbh=glm(rec~q2$DBH, family="binomial")
> summary(q2.dbh)
Call:
glm(formula = rec ~ q2$DBH, family = "binomial")
Deviance Residuals:
   Min     1Q  Median     3Q     Max
-3.6595  -1.6123  -0.8911  0.4493  7.4964
Coefficients:
        Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.829832  0.045977   -39.8  <2e-16 ***
q2$DBH      -0.026250  0.002345   -11.2  <2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)
   Null deviance: 1509.1  on 280  degrees of freedom
Residual deviance: 1286.0  on 279  degrees of freedom
AIC: 1626.1
Number of Fisher Scoring iterations: 5
```

4. GRAPH

```
> plot(rec[,1]~q2$DBH, ylab="proportion of successful recruits",
xlab= "amount of conspecific adult trees (DBH)")
> curve(plogis(-1.83+-0.03*x), add=T, col="red", lwd=2)
```
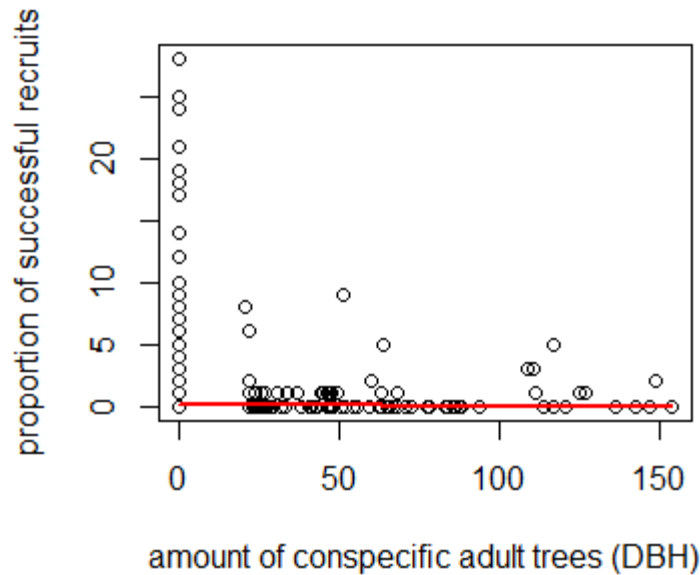
Fig.3. Effect of amount of conspecific adult trees in seed implantation survival. Proportion of successful seed implantation is -1.83 when number of conspecific trees is at 0. For every presence of adult conspecific tree, proportion of recruits decreases By 0.03 (*P<2e-16*). Red line shows line of best fit (AIC:1626.1).

- **Results suggest that presence of conspecific trees (adults) influence seed germination.**

Q3: MOSQUITO DATA

```
> str(q3)
'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 ...
```

1.PROPORTION OF SURVIVED MOSQUITOES

```
> surv=cbind(q3$Emergent_adults, q3$Egg_Count-q3$Emergent_adults)
> prop3=(q3$Emergent_adults/ q3$Egg_Count)
> p3=glm(prop3~q3$Detritus, family="binomial")
```

2.GRAPHS

```
> plot(prop3~q3$Detritus, ylab= "Emergent Adults (proportion)", xlab=" Detritus")
> coef(p3)
(Intercept) q3$Detritus
  1.2722258  -0.2975641
> curve(plogis(1.27+-0.30*x), add=T, col="red", lwd=3)
> #polynomial
> curve(plogis(1.44-.19*x+-0.21*x^2+0.04*x^3),add=T,col="blue",lwd=3)
> #Ricker
> curve(plogis(10*x*exp(-2*x)),col="green",add=T,lwd=3)
```
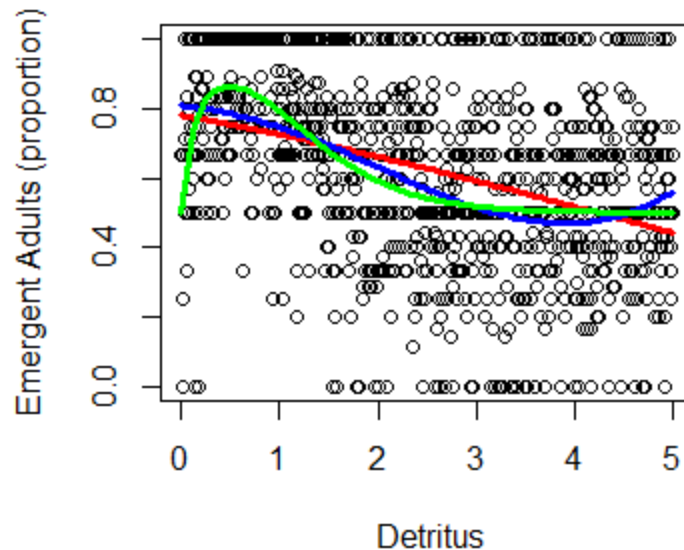
Fig.4. Effect of detritus on proportion of emergent adult mosquitoes. Red line corresponds to line of best fit. Blue line corresponds to polynomial equation of the line and green corresponds to Ricker equation of the line. Line of best fit suggests that proportion of emergent adults is at 1.27 and that for every 1 unit increase of detritus, Proportions are reduced by 0.30 (*P*= 2.73e-10; AIC: 1258.3).

- Polynomial equation suggests that proportion of emergent adults decrease with increasing detritus. However, beyond 5 units of detritus, proportions begin to increase again. For Ricker equation, it is suggested. That peak emergence of adults occurs at 1 unit of detritus and as detritus levels increase, the proportions decrease. Further, beyond 3 units of detritus, proportions remain constant.

```
> #Ricker
> -sum(dbinom(q3$Emergent_adults,prob=plogis(10*x*exp(-2*x)),size=length(q3$Egg_Co
unt),log=T))
[1] 993876.4
> #polynomial
> -sum(dbinom(q3$Emergent_adults,prob=plogis(1.44-.19*x+-0.21*x^2+0.04*x^3),size=l
ength(q3$Egg_Count),log=T))
[1] 963626.5
```

- The Ricker model has a higher likelihood than the polynomial model which means that the probability of adult mosquito emergence at a certain threshold. This may be backed up by the logistic growth model for survival.

Q4: DATA SIMULATION
A.LINEAR

```
> #effect of amount of DDT in soil to proportion of raptor deaths
> intercept=5
> slope=2
> n= 250
> ddt=seq(from=0, to=100, length=n)
> deaths=rnorm(n, mean=intercept+slope*ddt, sd=5)
> plot(deaths~ddt, ylab=" No.of raptor deaths", xlab="amount of DDT in the so
il (ppm)")
```
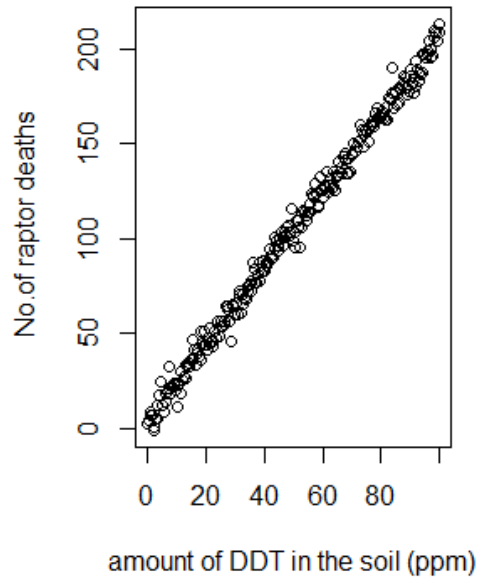
Fig.5. Effect of DDT on number of recorded raptor deaths. As amount of DDT in the soil increases so does no.of raptor deaths.

```
> samsize=rep(seq(from=10, to=250), times=20)
> estslope=rep(NA, times=length(samsize))
> #for loop
> for(i in 1: length(samsize)){
+   y=rnorm(n=samsize[i], mean=intercept+slope*seq(from=0, to=100, length=sam
size[i]), sd=5)
+    response=y
+    q4glm=glm(response~seq(from=0, to=100, length=samsize[i]))
+    estslope[i]=coef(q4glm)[2]
+ }
> plot(estslope~samsize)
```
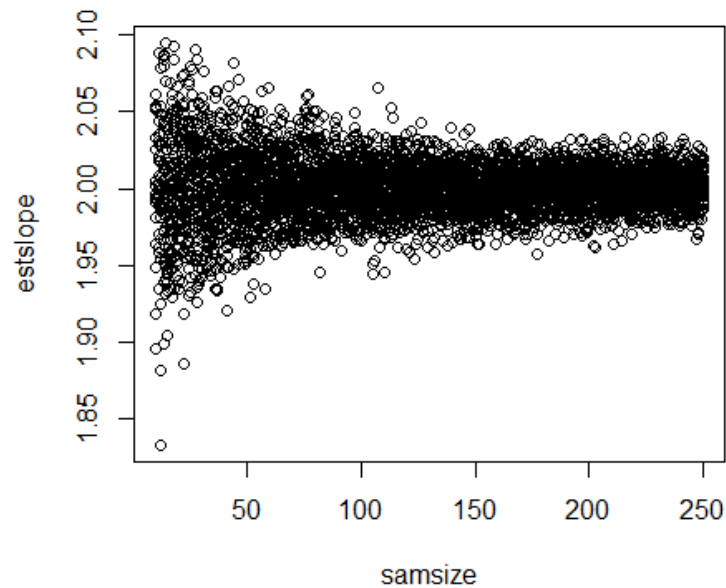
Fig.6. Power analysis on the sample size needed to detect significant effect of DDT on raptors. Results suggest that little variation from the estimated slope of 2 is observed when sample size is >150. Thus, records of >150 raptor deaths caused by DDT poisoning will be enough to prove its effect.

B.BINOMIAL

```
> int1=4
> sl1=-2
> n=1000
> hatch=rbinom(n=1000, prob=plogis(int1+sl1*amtddt), size=5)
> amtddt=seq(from=1, to=10, length=1000)
> plot(hatch~amtddt, ylab="no. of successful hatchlings", xlab=" amt. of DDT
in soil (ppm)")
```
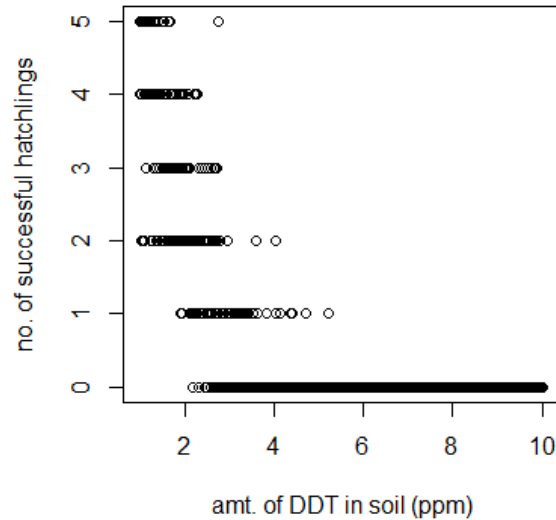
Fig.7. Effect of amt of DDT in the soil on hatchling success. At increasing levels (>2 ppm) of DDT de tected in the soil, success rate for hatchlings is at 0.

```
> #for loop:
> nsize=rep(seq(from=10, to=100), times=20)
> estimsl=rep(NA, times=length(nsize))
> for(k in 1: length(nsize)) {
+    yy=rbinom(n=nsize[k], prob=plogis(int1+sl1*seq(from=1, to=10, length=nsiz
e[k])), size=5)
+    respon=cbind(yy, 5-yy)
+    q4g=glm(respon~seq(from=1, to=10, length=nsize[k]), family="binomial")
+    estimsl[k]=coef(q4g)[2]
+ }
> plot(estimsl~nsize, ylab="estimated slope", xlab="sample size")
```
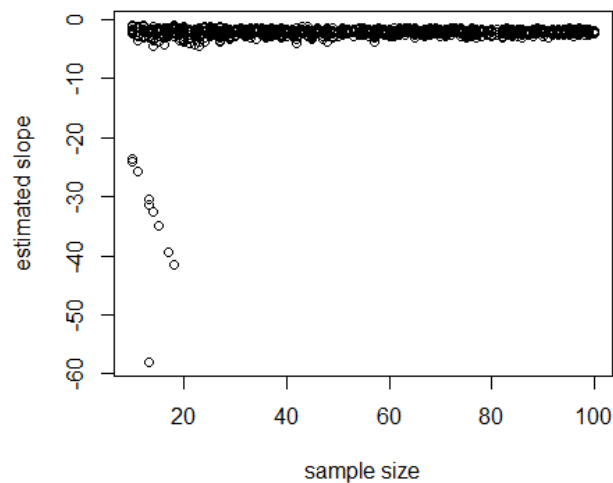

Fig.8. Power analysis on the sample size needed to detect significant effect of DDT on hatchling su ccess. Results suggest that little variation from the estimated slope of -2 is observed when sample size is >30. Thus, records of >30 hatchlings (egg hatches) will be enough to prove the effect of DDT on hatchling success (egg thinning).

Model	Sample size to accurately estimate slope	Slope if p<0.05	Sample size if Slope p<0.05 (df)
Binomial	30	-1.4720	54
Linear	150	1.99	249

In general, discrete data have less power than continuous data because the decimal points add precision, and thus holds more information than a binary counterpart. In effect, continuous data require fewer samples than discrete which is not reflective of my results possibly because I set the two equations very differently.