homework2

Trevor Caughlin

January 20, 2018

<https://ms.mcmaster.ca/~bolker/emdbook/book.pdf> Likelihood chapter beginning on page 228

homework 1: Run a logistic regression on these data; what is your interpretation of the slope? *make it a real data set* present the confidence interval

homework 2: run a binomial regression on these data

present the confidence intervals

1. use glm to analyze the dataset “SEEDLING\_SURVIVAL.csv” These data represent the annual survival rates of 1435 tagged tree seedlings in Huai Kha Khaeng Wildlife Sanctuary, Thailand. Run separate glm models for:
2. effect of height on seedling survival
3. effect of light on seedlings survival

As output, include a sentence with parameter interpretation for both models: “The baseline survival was…, The effect of height on seedling survival was …” include two separate plots with curve to overlay model predictions on data Also include a confindence intervals using confint function

From your results, which predictor variable has a stronger impact on output?

1. plot the data. Note that plotting the proportion of eggs is necessary, since the eggs and mosquito larvae both vary
2. Which is a better model to fit the data? polynomial

just do in word 1.44+-0.19*x+-0.21*x^2+0.04\*x^3

10*x*exp(-2\*x)

1. use curve with add=T to overlay curves with both parameters on the data
2. use dbinom to calculate the likelihood of both models (with given parameter values). Note: be sure to use log=T for taking sum of dbinom output

or some other function? use dbinom, log and sum

homework 3: power analysis use rbinom and rnorm glm for both

note this will involve a for loop

step 1: create an empty vector for for loop to fill in Here, the object empty vector is just NA repeated 100 times. We are going to fill in these values using the for loop.

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

In this example, i is a counter for each iteration of the for loop (from 1 to 100). At each of i iterations, R is filling in empty\_vector with a draw from rnorm

how many samples do you need to (a) recapture the slope value?

1. ensure a p-value for slope parameter <0.05

output should include: a) numeric values b) plot of sample size (x-axis) and power (either p-value or bias in parameters) on y-axis