HW5

Louis Jochems

3/11/2019

##Question 1##  
#a)   
library("rstanarm")

## Loading required package: Rcpp

## rstanarm (Version 2.15.2, packaged: 2017-04-19 09:04:44 UTC)

## - Do not expect the default priors to remain the same in future rstanarm versions.

## Thus, R scripts should specify priors explicitly, even if they are just the defaults.

## - For execution on a local, multicore CPU with excess RAM we recommend calling

## options(mc.cores = parallel::detectCores())

#b) https://discourse.mc-stan.org/t/multiple-outcomes-hurdle-model-exists/7951  
  
# I would want apply my EFB abundance/presence absence model into a hurdle model with probabilistic intepretations  
# compared to a hurdle model intrepreting the conventional parameters to see what new information I may learn. In particular,   
#this forum discusses how to deal with latent (random) variables that may account for a good deal of covariance in a  
# large data set.

##Question 2##

#Question 3  
#posterior probablility of random pixel being land pixel from earth =0.23   
#equally likely pixel from earth or mars  
prob=0.5  
#earth is 30% land  
EarthLand=0.3   
EarthWater=0.7  
#mars is 100% land   
MarsLand=1  
  
postprob=(EarthLand\*prob)/((MarsLand+EarthLand)/2)   
#EarthLand is probability of our data given our hypothesis, prob is prior probability that pixel equally being from earth or mars, denominator represents probability of finding land pixel from both planets (average likelihood)  
  
#ANSWER: 23% chance that sampled pixel will be Earth land

##Question 4##  
#2 spp panda, equally common, look alike, no genetic assay   
prob=0.5  
#Sp A birth outcomes  
BirthA\_twins=0.1  
BirthA\_infant=0.9   
#Sp B birth outcomes  
BirthB\_twins=0.2  
BirthB\_infant=0.8  
  
#A) Probability that panda is sp. A, twins @ first birth. BirthA\_twins is our data given our hypothesis, prob is prior that both panda spp. are equally common denominatior is likelihood of having twins in general for both spp (avg likelihood).   
post\_prob=BirthA\_twins\*prob/((BirthA\_twins+BirthB\_twins)/2)   
  
#33% chance panda is sp. A and gives first birth to twins   
  
  
#B)Probability that panda is sp. A, second birth is infant after giving birth to twins. Twins then infant represents our data given conjecture (in that order) #prob is our prior that both sp. are equally common. Denominator represents likelihood conjecture sequence for both species   
post\_probATI=(BirthA\_twins\*BirthA\_infant\*prob)/(((BirthA\_twins\*BirthA\_infant)+(BirthB\_twins\*BirthB\_infant))/2)  
  
#36% chance that this panda is spp. A after giving first birth to twins and second birth to single infant. Slight increase makes sense given that sp. A has slightly higher chance of giving birth to singleton infant.   
  
# probability that next birth will also be twins   
post\_prob=(BirthA\_twins\*BirthA\_twins\*prob)/(((BirthA\_twins\*BirthA\_infant)+(BirthB\_twins\*BirthB\_infant))/2)  
#There's only a 4% chance that next birth will also be twins.   
  
#C) new genetic test comes along and our panda test positive for spp A   
#prob that test is correct for spp A   
probAA=0.8  
#prob that test incorrectly identifies A but is actually B  
probBA=0.2  
  
#prob that test is correct for spp B  
probBB=0.65  
#prob that test incorrecctly identifies B but is actually A  
probAB=0.35  
  
###genetic test w/o birth data. probAA is probability of our data given our conjecture for that test is correct in showing A. prob still represents equal likelihood given   
#that both species are equally common. Denominator represents likelihood for test being A.###  
postgenA=probAA\*prob/((probAA+probAB)/2)  
  
#69.5% chance that genetic test correctly identifies panda as sp. A.   
  
  
###post prob of genetic test w/ birth data. genAbirth represents prob that test correctly id's sp A (either true positive or false neg) with the prior being   
#set as the posterior probability of birth data conjecture (Sp A, twins, infant) + prob incorrectly id's sp A (false negative) with prior's probability  
#that birth data was not sp A with twins then infant. #AA test represents posterior probaility of test correctly IDing sp A with birth data.   
  
genABirth=probAA\*post\_probATI+probAB\*(1-post\_probATI)  
AAtest=(probAA\*post\_probATI)/genABirth  
   
#Test probability is now only 56% chance that the tested panda is from species A given the previous data on births.