Daniella Assing

Geosci 541

April 25 2016

Lab 13

Problem Set 1

There is a longstanding story that Triassic Diapsids outcompeted Triassic Syanpsids. Let's see if Triassic Diapsids were more likely to survive the Traissic/Jurassic extinction than Synapsids.

Question 1

Download four data sets from the paleobiology database. First, a dataset of Anisian-Rhaetian Synapsids, name it TriassicSynapsids. Second, a dataset of Anisian-Rhaetian Diapsids, name it TriassicDiapsids. Third, a dataset of post-Triassic Diapsids, name it JurassicDiapsids. Fourth, a dataset of post-Triassic Synapsids, name it JurassicSynapsids. Show your code.

TriassicSynapsids<-downloadPBDB("Synapsida","Anisian","Rhaetian")  
TriassicDiapsids<-downloadPBDB("Diapsida","Anisian","Rhaetian")  
JurassicDiapsids<-downloadPBDB("Diapsida","Jurassic","Neogene")  
JurassicSynapsids<-downloadPBDB("Synapsida","Jurassic","Neogene")

Question 2

How many Diapsid genera were there in the Triassic dataset? How many Synapsid genera? Show your code.

#clean  
TriassicSynapsids<-cleanRank(TriassicSynapsids,"genus")  
TriassicDiapsids<-cleanRank(TriassicDiapsids,"genus")  
JurassicDiapsids<-cleanRank(JurassicDiapsids,"genus")  
JurassicSynapsids<-cleanRank(JurassicSynapsids,"genus")

#number of Triassic diapsid genera

> length(unique(TriassicDiapsids[,"genus"]))  
[1] 388

#Trassic synapsid genera

> length(unique(TriassicSynapsids[,"genus"]))  
[1] 116

Question 3

How many Triassic Diapsid genera survived the Triassic/Jurassic transition? How many were victiims? How many Triassic Synapsid genera surivived the Triassic/Jurassic Transition? How many were victims? Show your code.

>#SURVIVORS

> TriassicDiapsidSurvivor<-intersect(TriassicDiapsids[,"genus"],unique(JurassicDiapsids[,"genus"]))  
> length(TriassicDiapsidSurvivor)  
[1] 38

> TriassicSynapsidSurvivor<-intersect(TriassicSynapsids[,"genus"],unique(JurassicSynapsids[,"genus"]))  
> length(TriassicSynapsidSurvivor)  
[1] 9

>#VICTIMS

> TriassicDiapsidVictim<-setdiff(TriassicDiapsids[,"genus"],unique(JurassicDiapsids[,"genus"]))

> length(TriassicDiapsidVictim)

[1] 350

>

> TriassicSynapsidVictim<-setdiff(TriassicSynapsids[,"genus"],unique(JurassicSynapsids[,"genus"]))

> length(TriassicSynapsidVictim)

[1] 107

Question 4

Calculate the odds ratio and log-odds that Diapsid genera were more likely to survive the T/J transition than Synapsids.

#DIAPSID ODDS

>DiapsidSurvivalOdds<-(length(TriassicDiapsidSurvivor)/length(TriassicDiapsids))/(length(TriassicDiapsidVictim)/length(TriassicDiapsids))

> DiapsidSurvivalOdds

[1] 0.1085714

#SYNAPSID ODDS

>SynapsidSurvivalOdds<-(length(TriassicSynapsidSurvivor)/length(TriassicSynapsids))/(length(TriassicSynapsidVictim)/length(TriassicSynapsids))  
> SynapsidSurvivalOdds  
[1] 0.08411215

#ODDS RATIO

> OddsRatio<-DiapsidSurvivalOdds/SynapsidSurvivalOdds  
> OddsRatio  
[1] 1.290794

#LOG-ODDS

> log(OddsRatio)  
[1] 0.2552573

Question 5

Using a 95% confidence interval, can you say that this odds/ratio is "statistically significant"? Show your code.

> StandardError<-sqrt(1/length(TriassicDiapsidSurvivor) + 1/length(TriassicSynapsidSurvivor)+1/length(TriassicDiapsidVictim)+1/length(TriassicSynapsidVictim))

>

> StandardError

[1] 0.3868202

#95% CONFIDENCE LIMIT

#UPPER LIMIT

> UpperLimit<-log(OddsRatio)+(StandardError\*1.96)  
> UpperLimit  
[1] 1.013425

#LOWER LIMIT  
> LowerLimit<-log(OddsRatio)-(StandardError\*1.96)  
> LowerLimit  
[1] -0.5029103

Since the lower limit is a negative number, the odds-ratio is not statistically significant.

Problem Set 2

Let's apply the technique that you just learned the Triassic and Jurassic Diapsids and Synapsids.

Queston 1

Download a dataset of Anisian-Rhaetian Diapsids and Synapsids, and a dataset of post-Triassic Diapsids and Synapsids. Show your code.

> TriassicTaxa<-downloadPBDB(c("Diapsida","Synapsida"),"Anisian","Rhaetian")  
> TriassicTaxa<-cleanRank(TriassicTaxa,"genus")  
>   
> JurassicTaxa<-downloadPBDB(c("Diapsida","Synapsida"),"Jurassic","Neogene")  
> JurassicTaxa<-cleanRank(JurassicTaxa,"genus")

Question 2

Find the mean latitude of each genus's occurrences in your Triassic dataset. Show your code.

> MeanLatitudes<-tapply(TriassicTaxa[,"paleolat"],TriassicTaxa[,"genus"],mean)

> head(MeanLatitudes)  
 Acaenasuchus Acallosuchus Acompsosaurus Actiosaurus Adamanasuchus Adelobasileus   
 10.116 10.430 10.740 32.120 10.145 10.170

Question 3

Find which Triassic genera were survivors and which were victims of the Triassic/Jurassic event. Show your code.

#TRIASSIC SURVIVORS

>TriassicSurvivors<-subset(TriassicTaxa,TriassicTaxa[,"genus"]%in%unique(JurassicTaxa[,"genus"])==TRUE)  
> TriassicSurvivors<-unique(TriassicSurvivors[,"genus"])  
> head(TriassicSurvivors)  
[1] "Clevosaurus" "Grallator" "Rhynchosauroides" "Rotodactylus"   
[5] "Brachychirotherium" "Coelurosaurichnus"

#TRIASSIC VICTIMS

>TriassicVictims<-subset(TriassicTaxa,TriassicTaxa[,"genus"]%in%unique(JurassicTaxa[,"genus"])!=TRUE)  
> TriassicVictims<-unique(TriassicVictims[,"genus"])  
> head(TriassicVictims)  
[1] "Icarosaurus" "Rutiodon" "Kuehneosuchus" "Kuehneosaurus"   
[5] "Trilophosaurus" "Diphydontosaurus"

Question 4

Find which genera of your Triassic dataset were Diapsids and which were Synapsids. Show your code.

>OnlyTriDiapsids<-subset(TriassicTaxa,TriassicTaxa[,"genus"]%in%TriassicDiapsids[,"genus"]==TRUE)

>OnlyTriDiapsids<-unique(OnlyTriDiapsids[,"genus"])

>OnlyTriSynapsids<-subset(TriassicTaxa,TriassicTaxa[,"genus"]%in%TriassicSynapsids[,"genus"]==TRUE)

>OnlyTriSynapsids<-unique(OnlyTriSynapsids[,"genus"])

Question 5

Perform a logistic regression where the outcome variable is Survivor/Victim and the input variable is the mean latitude of each genus. Show your code. Was the mean latitude of a Triassic genus a good predictor of its survival across the T/J extinction?

> PTVictims<-array(0,dim=length(TriassicVictims),dimnames=list(TriassicVictims))  
> head(PTVictims)  
 Icarosaurus Rutiodon Kuehneosuchus Kuehneosaurus Trilophosaurus   
 0 0 0 0 0   
Diphydontosaurus   
 0

> FinalMatrix<-merge(MeanLatitudes,PTVictims,all=TRUE,by="row.names")  
> head(FinalMatrix)  
 Row.names x y  
1 Acaenasuchus 10.116 0  
2 Acallosuchus 10.430 0  
3 Acompsosaurus 10.740 0  
4 Actiosaurus 32.120 0  
5 Adamanasuchus 10.145 0  
6 Adelobasileus 10.170 0

> FinalMatrix<-transform(FinalMatrix,row.names=Row.names,Row.names=NULL)  
> colnames(FinalMatrix)<-c("MeanLatitudes","Survivor/Victim")  
>   
> head(FinalMatrix)  
 MeanLatitudes Survivor/Victim  
Acaenasuchus 10.116 0  
Acallosuchus 10.430 0  
Acompsosaurus 10.740 0  
Actiosaurus 32.120 0  
Adamanasuchus 10.145 0  
Adelobasileus 10.170 0

> FinalMatrix[is.na(FinalMatrix[,"Survivor/Victim"]),]<-1  
>   
> head(FinalMatrix)  
 MeanLatitudes Survivor/Victim  
Acaenasuchus 10.116 0  
Acallosuchus 10.430 0  
Acompsosaurus 10.740 0  
Actiosaurus 32.120 0  
Adamanasuchus 10.145 0  
Adelobasileus 10.170 0

#PERFORM REGRESSION

> Regression<-glm(FinalMatrix[,"Survivor/Victim"]~FinalMatrix[,"MeanLatitudes"],family="binomial")  
> summary(Regression)  
  
Call:  
glm(formula = FinalMatrix[, "Survivor/Victim"] ~ FinalMatrix[,   
 "MeanLatitudes"], family = "binomial")  
  
Deviance Residuals:   
 Min 1Q Median 3Q Max   
-0.4529 -0.4466 -0.4440 -0.4362 2.1782   
  
Coefficients:  
 Estimate Std. Error z value Pr(>|z|)   
(Intercept) -2.2750384 0.1532604 -14.84 <2e-16 \*\*\*  
FinalMatrix[, "MeanLatitudes"] 0.0007675 0.0051136 0.15 0.881   
---  
Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
(Dispersion parameter for binomial family taken to be 1)  
  
 Null deviance: 312.48 on 503 degrees of freedom  
Residual deviance: 312.46 on 502 degrees of freedom  
AIC: 316.46  
  
Number of Fisher Scoring iterations: 5

Is latitude a good predictor?

The estimate is 0.0007675. This means that for every one degree increase in latitude, the odds of surviving the P/T increases by 0.0007675. Since this is such a small number, no, I don’t think that latitude is a good predictor of survival.

**Extra Credit (6 Points)**

Perform a *multiple* logistic regression where the outcome varaible is Survivor/Victim status and the input variables are the mean latitude of each genus *and* whether the gneus is a Diapsid/Synapsid. Is status as a Synapsid/Diapsid more or less important average paleolatitude of occurrences for survival? Show your code.

Hint:

* The general formula for a *multiple* logistic regression is: glm(Outcome ~ Variable1 + Variable2,family="binomial")
* You'll want to represent Diapsids and Synapsids with 1s and 0s, similarly to how we did survivor and victim status

Is status as a Synapsid/Diapsid more or less important average paleolatitude of occurrences for survival?

> TriassicArray<-array(1,dim=length(OnlyTriDiapsids),dimnames=list(OnlyTriDiapsids))

> Intermediate<-merge(MeanLatitudes,TriassicArray,all=TRUE,by="row.names")

> Intermediate<-transform(Intermediate,row.names=Row.names,Row.names=NULL)

> Final<-merge(Intermediate,TriassicArray,all=TRUE,by="row.names")

> Final<-transform(Final,row.names=Row.names,Row.names=NULL)

> colnames(Final)<-c("MeanLatitudes","Diapsids1/Synapsids0","Survivor/Victim")

> Final[is.na(Final[,"Diapsids1/Synapsids0"]),]<-0

> Final[is.na(Final[,"Survivor/Victim"]),]<-1

> Regression<-glm(Final[,"Survivor/Victim"]~Final[,"MeanLatitudes"] + Final[,"Diapsids1/Synapsids0"],family="binomial")

Warning message:

glm.fit: algorithm did not converge

> summary(Regression)

Nope, neither one is better than the other in my opinion because both of them have almost equal estimates (5.8 and 5.3). I don’t know what the warning message was. I looked it up online and found that I might want to use maxfit=100+. It did take away the error message and it also changed the probability (all probabilities=1) but it didn’t change the estimate results.

This was such a long lab..! D: