Final Project - EEB590C

Devin Molnau and Elizabeth McMurchie

April 29, 2021

Homework 3

This assignment is due prior to the last day class. You are to self-select and work in groups: 2-3 in a group. For the assignment below submit one R-script. Annotations via comments are highly encouraged. The script should run!

Assignment:

1: Obtain a dataset. This may be one of your own, a dataset from DRYAD, or some other source. Identify hypotheses for this dataset and analyze patterns in the data. You may use any methods learned during the semester, but at least one analysis must come from material learned in weeks 11-13.

USE COMMENTS IN THE R CODE to describe what the patterns you find represent.

```
setwd("D:/Documents/BoxSync/Classes_Spring2021/Advanced_Biostatistics_EEB590C/Homework/Final_EEB590")
#load appropriate libraries
library(knitr)
library(RRPP)
library(geomorph)
library(tidyverse)
library(readxl)
library(ade4)
library(vegan)
library(mice)
```

Intro

Epidermal micromorphology in Neotropical bamboo foliage leaves has proven to be a useful tool for differentiating between different species. Certain unusual features have been noted in the epidermal micromorphology of one species of savanna bamboo, Guadua paniculata, including papillae on both leaf surfaces and saddle-shaped silica cells. However, it remains unknown whether the few other species of Guadua primarily found in forests, including G. virgata, have similar micromorphology. Additionally, there are several general trends in Guadua macromorphology, including the G. angustifolia type, which are tall, primarily forest-adapted species, the G. glomerata type, which are scandent species primarily found along rivers, the G. sarcocarpa type, which are similar to G. angustifolia but notable for their tree-killing ability, and species of both savanna and forest bamboos that closely resemble G. paniculata. Here, we analyzed whether micromorphological features such as shape of silica bodies and presence, placement, and shape of stomata and papillae on epidermal cells of foliage leaves were associated with different patterns in macromorphology, habitat type, and country of origin of specimens belonging to a selection of species of Guadua.

Preprocessing

Read in data

```r
#READ IN DATA AS DATAFRAME
mydata<-read_excel(path="data/TransformGuaduaSet.xlsx", col_names = TRUE, na="x")
```

Remove Column V and Y

Remove column V and Y, because there is missing data and we cannot impute the data because it just doesn't make sense to attempt to predict these columns.

V. Adaxial: Frequency of stomates if present on the adaxial surface of foliage leaf blades: 0 = common; 1 = infrequent. Y. Adaxial: Papillae on long cells of the intercostal zone adjacent to the stomates: 0 = not overarching the stomates; 1 = overarching the stomates.

```r
to_drop<-c("Adaxial: Papillae on long cells of the intercostal zone adjacent to the stomates:  0 = not 

df=data.frame(mydata[,!( names(mydata) %in% to_drop)],stringsAsFactors = TRUE)
```

Set columns to factors

```r
df<-data.frame(lapply(df,as.factor))
```

Imputation of missing data

In the last three columns, in some species the papillae of the cells adjacent to the stomata obscured theshape of the subsidiary cells. Therefore, the shapes of these subsidiary cells resulted in missing data that we imputed using the MICE package's logistic regression method.

```r
init = mice(df,maxit=0)
meth = init$method
predM=init$predictorMatrix
meth[c(colnames(df[,7:length(df)]))]<-"logreg"
set.seed(183)
imputed<-mice(df,method = meth, predictorMatrix = predM, m=5,printFlag = FALSE)
```

```
## Warning: Number of logged events: 90
```

```r
imputed<-complete(imputed)
```

```r
sapply(imputed,function(x) sum(is.na(x))) #double check to make sure there are no more NA
```

```
##
##
##
##
##
##
##
##
##
##
##
##
##                                Abaxial.papillae.on.the.long.cells.in.the.stomatal.zone.0...absent.1...
##
##                      Abaxial.papillae.on.the.long.cells.in.the.interstomatal.zone...0...absent..1...p
##
##                                                                        Abaxial..Prickle.hairs..0...absent..1...
```

```
##
##                                                    Abaxial..Bicellular.microhairs..0...absent..1...
##
##                                                          Abaxial..Macrohairs..0...absent..1...
##
##                        Abaxial..Dumbbell.shaped.silica.bodies.in.the.intercostal.zone...0...present..1..
##
##                         Abaxial..Saddle.shaped.silica.bodies.in.the.intercostal.zone...0...present..1..
##
##                   Abaxial..Crenate..elongated..silica.bodies.in.the.intercostal.zone...0...present..1...
##
##                              Abaxial..Dumbbell.shaped.silica.cells.in.the.costal.zone...0...present..1...
##
##                               Abaxial..Saddle.shaped.silica.cells.in.the.costal.zone...0...present..1...
##
##                         Abaxial..Crenate..elongated..silica.cells.in.the.costal.zone...0...present..1...
##
##                                          Abaxial..Triangular.stomata.subsidiary.cells..0...absent..1...
##
##                                        Abaxial..Dome.shaped.stomata.subsidiary.cells..0...absent..1...
##
##                                    Abaxial..Parallel.sided.stomata.subsidiary.cells..0...absent..1...
##
##                      Adaxial..stomates.on.the.adaxial.surface.of.foliage.leaf.blades...0...absent..1...pr
##
## Adaxial.papillae.on.the.long.cells.in.the.stomatal.zone.exclusive.of.bulliform.cells.0...absent.1...
##
##                           Adaxial.papillae.on.the.long.cells.in.the.interstomatal.zone...0...absent..1...pr
##
##                                         Adaxial.papillae.on.the.bulliform.cells...0...absent..1...pr
##
##                                                          Adaxial..Prickle.hairs..0...absent..1...
##
##                                                    Adaxial..Bicellular.microhairs..0...absent..1...
##
##                                                          Adaxial..Macrohairs..0...absent..1...
##
##                        Adaxial..Dumbbell.shaped.silica.bodies.in.the.intercostal.zone...0...present..1..
##
##                         Adaxial..Saddle.shaped.silica.bodies.in.the.intercostal.zone...0...present..1..
##
##                   Adaxial..Crenate..elongated..silica.bodies.in.the.intercostal.zone...0...present..1...
##
##                              Adaxial..Dumbbell.shaped.silica.cells.in.the.costal.zone...0...present..1...
##
##                               Adaxial..Saddle.shaped.silica.cells.in.the.costal.zone...0...present..1...
##
##                         Adaxial..Crenate..elongated..silica.cells.in.the.costal.zone...0...present..1...
##
##                                          Adaxial..Triangular.stomata.subsidiary.cells..0...absent..1...
##
##                                        Adaxial..Dome.shaped.stomata.subsidiary.cells..0...absent..1...
##
##                                    Adaxial..Parallel.sided.stomata.subsidiary.cells..0...absent..1...
```

##

Analysis

PCOA

```
#TODO
#dist.binary() #Distance matrix for binary is simple matching coefficient
```

STICK PLOT of PCOA

```
#TODO
#example
#screeplot(pca.bumpus,bstick = TRUE)
```

FACTORIAL ANOVA

```
#TODO
```

INTERACTIONS

```
#TODO
#### Group VS country
#### Group vs habitat
```

MODEL COMPARISON USING LIKELIHOOD RATIO TEST (LTR)

```
#TODO
```